JUNE
1954
VOLUME 18
NUMBER 3

JOURNAL OF CONSULTING PSYCHOLOGY

AMERICAN PSYCHOLOGICAL ASSOCIATION

JOURNAL OF CONSULTING PSYCHOLOGY

Edited by Laurance F. Shaffer

Teachers College, Columbia University

Associate Editors: Edgar A. Doll, *Bellingham, Wash.* Ann Magaret Garner, *University of Illinois.* William A. Hunt, *Northwestern University.* E. Lowell Kelly, *University of Michigan.* Morris Krugman, *Board of Education, New York, N. Y.* Bertha M. Luckey, *Board of Education, Cleveland, Ohio.* Fred McKinney, *University of Missouri.* Anne Roe, *New York, N. Y.* Carl R. Rogers, *University of Chicago.* R. Nevitt Sanford, *University of California.* Edward Joseph Shoben, Jr., *Teachers College, Columbia University.* J. R. Wittenborn, *Yale University.*

Lorraine Bouthilet, *Managing Editor*

The *Journal of Consulting Psychology* is the clinical journal of the American Psychological Association. It is devoted primarily to the publication of original research relevant to psychological diagnosis, psychotherapy, and counseling, personality, and the dynamics of behavior. Case studies, theoretical contributions, descriptions of clinical techniques, and discussions of the professional problems of clinical psychology also appear in its pages.

Editorial Office

Manuscripts submitted for publication, books and tests for review, and correspondence on editorial matters should be sent to the Editor, Laurance F. Shaffer, 525 West 120th St., New York 27, N. Y.

Information for contributors: The important characteristics of a good article are the significance of its subject matter, the demonstrated soundness of its conclusions, its clarity, and its brevity. Clarity is improved by careful outlining, discriminating use of headings, and a simple, straightforward style of writing. For brevity, authors should strive to convey every thought without waste of words, and should avoid repetitions of methods, data or findings.

The style of all manuscripts should conform to the instructions given in the Publication Manual of the American Psychological Association, prepared by the Council of Editors (*Psychological Bulletin*, 1952, 48, 389-449). If it is not otherwise available, the *Manual* may be purchased, for $1.00, from the Washington office of the Association. The Editor expects each author to study the *Manual* carefully before submitting a manuscript. Special attention should be given to the style of headings, capitalization, indention, spacing, and footnotes, and to the preparation of tables, figures, and lists of references. Otherwise acceptable manuscripts which depart seriously from the standards of the *Manual* are returned to the author for revision and retyping.

There is a lag of about nine to twelve months between the acceptance of an article and its publication. Authors may secure immediate publication, however, by arranging for articles to be printed as extra pages, increasing the number of pages that subscribers receive. The cost to the author of such early publication is approximately $20.00 per printed page.

Authors receive fifty reprints gratis, and may order additional quantities when returning proof.

Business Office

Communications concerning subscriptions, change of address, claims for the nonreceipt of a number, advertising, and other business matters, should be sent to the American Psychological Association, Inc., 1333 Sixteenth St., N. W., Washington 6, D. C.

The Journal of Consulting Psychology is published bimonthly in February, April, June, August, October, and December, the yearly volume comprising approximately 480 pages. The subscription per year is $7.00, foreign $7.50, single number $1.25.

Copyright, 1954 by the American Psychological Association, Inc. Printed and mailed by the Dentan Printing Co., 23 West Colorado Ave., Colorado Springs, Colo. and entered as second-class matter at the Post Office at Colorado Springs, Colo. Acceptance for mailing at special rate of postage provided for in section 538, Act of February 25, 1925, authorized May 27, 1940.

JOURNAL OF CONSULTING PSYCHOLOGY

June, 1954 Vol. 18, No. 3

CONTENTS

Prognostic Factors in Schizophrenia: William *Schofield, Starke R. Hathaway, Donald W. Hastings, and Dorothy M. Bell* - - - - - - - - - 155

Direction of Aggression as a Prognostic Variable in Mental Illness: *Dorothy A. Feldman, G. R. Pascal, and C. H. Swensen* - - - - - - - - - 167

Prognostic Significance of Precipitating Stress in Mental Illness: *Mary E. Cole, C. H. Swensen, and G. R. Pascal* - - - - - - - - - - - 171

Patient Cancellations Following a Sudden Cancellation by the Therapist: *Miriam Crowley McCue* - - - - - - - - - - - - - - - - - - - 176

The Comparative Effectiveness of Individual Psychotherapy, Group Psychotherapy, and a Combination of These Methods: *George O. Baehr* - - - - - - - - - 179

Body Size and Body-Cathexis: *Sidney M. Jourard and Paul F. Secord* - - - - - - 184

Validation of Rorschach Criteria for Predicting Duration of Therapy: *Robert G. Gibby, Bernard A. Stotsky, E. Wesley Hiler, and Daniel R. Miller* - 185

A Rationally Derived MMPI Scale to Measure Dependence: *Leslie Navran* - - - - 192

Comparison of the Rorschach and MMPI by Means of Factor Analysis: *Harold L. Williams and James F. Lawrence* - - - - - - - - - - - - - 193

The Diagnosis of Subclinical Schizophrenia: *Donald R. Peterson* - - - - - - - - 198

The Relationship Between Anxiety Level and Several Measures of Intelligence: *Joseph D. Matarazzo, George A. Ulett, Samuel B. Guze, and George Saslow* - - 201

Comparison of Two Forms of the Manifest Anxiety Scale: *Joyce B. McCreary and A. W. Bendig* - - - - - - - - - - - - - - - - - 206

Rater Experience and the Reliability of Case History Ratings of Adjustment: *A. W. Bendig and Janine Sprague* - - - - - - - - - - - - - - - - - - 207

Personality Correlates of Undergraduate Field of Specialization: *Richard C. Teevan* - 212

Congruence of Self and Ideal Self in Relation to Personality Adjustment: *Thomas E. Hanlon, Peter R. Hofstaetter, and James P. O'Connor* - - - - - - 215

Validity of a Critical-Item Scale for Schizophrenia on the MMPI: *Harold Rubin* - - 219

Wechsler-Bellevue Performance of Delinquent and Nondelinquent Girls: *Julia R. Vane and Virginia W. Eisen* - - - - - - - - - - - - - - - - 221

Note on Sanders and Pacht's Investigation of Size Constancy of Clinical Groups: *Edward Lovinger* - - - - - - - - - - - - - - - - - - - 226

New Books and Tests - 229

Brief Reports:

A New Publication Procedure

For a trial period, the *Journal of Consulting Psychology* will accept *"brief reports"* for early publication without expense to the author, under certain conditions. The new procedure is intended to reduce publication lag and permit the circulation of worth-while studies of specialized interest or minor importance which cannot be accepted at present because of limitations of space.

A *brief report,* strictly limited to one printed page, may be accepted if the author agrees to make a more extended report available. Up to four pages in each issue may be devoted to brief reports, published in the order of their receipt without regard to the dates of receipt of articles accepted for regular publication. It is anticipated that most of the brief reports can appear in the first or second issue to go to press following their acceptance.

The procedure for submitting a *brief report* is:

1. The author sends the editor *both* the brief report and a more complete report of the study. The full report should be of sufficient length to give a clear account of the background, procedure, results, and conclusions of the study.

2. The *full report* is prepared in the style specified by the *Publication Manual* [1], except that it may be typed with single spacing. The editor will send the full report to the ADI.

3. The *brief report* gives a clear, condensed summary of the procedure of the study, and as full an account of the results as space permits. To insure that the brief report does not exceed one printed page, the author prepares it according to these specifications:

The text of the brief report, including all matter except the title and the author's lines, does not exceed 80 typewritten lines averaging 42 characters and spaces in length. Set the typewriter margins for short lines of 42 characters, which are 3.5 inches long in elite typing, or 4.2 inches long in pica. The manuscript is double spaced throughout and, except for the short lines, follows the standard style [1].

Headings, tables, and references are avoided in the brief report. If essential, they must be counted in the 80 lines. The brief report includes, within its 80 lines, a footnote in this style:[1]

[1]An extended report of this study may be obtained without charge from John Doe, 300 Market St., Prospect 6, Mass. [giving the author's full name and address] or for a fee from the American Documentation Institute. To obtain it from the latter source, order Document No. .... from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting $..... for microfilm or $..... for photocopies.

4. The author of a brief report prepares at least 100 mimeographed copies of the extended report. He agrees: (*a*) to send copies of the extended report upon request as long as the supply lasts, and (*b*) not to submit the extended report for publication to another printed journal of general circulation.

Reference

1. American Psychological Association. Council of Editors. Publication manual of the American Psychological Association. *Psychol. Bull.*, 1952, 49, 389-449.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Prognostic Factors in Schizophrenia[1]

William Schofield, Starke R. Hathaway, Donald W. Hastings, and Dorothy M. Bell
University of Minnesota

In view of the magnitude of the problem presented by the psychiatric enigma known as schizophrenia, it is startling to find how relatively little careful attention has been paid to the question of outcomes in this illness and how few comprehensive investigations have been made into prognostic factors. A reasonably exhaustive review of the research literature on outcome and prognosis in schizophrenia reveals only 31 published studies in this country between 1909 and 1951 — less than one report per year. These 31 studies cover a total of only 6,802 patients (exclusive of Pollock's [7] statistical review of 19, 927 patients), the different samples ranging in size from 20 to 1,082 patients. The length of the follow-up interval has ranged from zero to fifteen years, with an average of approximately four and one-half years in those studies which clearly specify the length of follow-up. Of the 31 studies reported, fewer than one-third have used direct face-to-face contacts with the patients as a basis for the follow-up evaluations, the remainder relying on questionnaire responses and letters. In view of the extent of the variability in the factors of sample size, length of follow-up period, and basis of follow-up evaluation, it is not surprising that the figures reported for "recovery" rate range from 0 to 35%, and that those for "improvement" rate extend from 5 to 74%!

A survey of the literature reveals that true investigations into prognosis, i.e., attempts to ascertain factors differentiating patients with good outcomes from those with poor, constitute a very small proportion of the follow-up studies, the majority of which are simple reports of the status of a given sample of patients at some period following hospital admission or discharge. The worker wishing to review specifically prognostic studies labors under some difficulty because of the fairly widespread tendency of investigators to use the term "prognosis" in studies which are purely statistical reports on "outcomes," with no effort to analyze factors related to outcome.

No attempt will be made in this paper to present a detailed review of the literature summarized above. A comprehensive review of the published studies on outcome and prognosis in schizophrenia, with a detailed bibliography, has been prepared for separate publication; the European researches are included in this review. However, it is appropriate here to call attention to a few papers which have particular import. In 1938, Romano and Ebaugh [8] published one of the earliest of the few studies in which both questionnaire responses and the results of direct re-examination were available for the same sample of patients as bases on which to appraise their postdischarge status. Comparison of the results of direct examination with questionnaire data from the families showed significant error in over 50% of the cases for which both types of data were available.

The paper by Malamud and Render [6] is important on a number of counts. It is one of the few studies that recognizes the marked discrepancies among the "recovery and improvement" statistics previously reported, and it reviews the possible sources of these discrepancies. It is concerned with an actual attempt to distinguish between schizophrenics with poor and good outcomes, respectively. It is based on a sample of over 300 patients and evaluates their status after an average postadmission interval of six years. Most important of all, the follow-up evaluation was

[1]This study was made possible by the Frederick B. Wells, Jr. Trust Fund. A condensation of this paper was read at the meeting of the American Psychological Association, Washington, September, 1952.

based on direct re-examination; and all patients had been originally examined by one of the investigators.

In 1941, Chase and Silverman [3] reported a "critical survey of the literature" on prognostic criteria in schizophrenia. This was by no means a comprehensive review, since the reference list of 49 items failed to include 15 of the 30 follow-up studies published in the United States through 1940. Thus, the degree of agreement or disagreement as to the prognostic value of particular items might actually prove quite different from that indicated by Chase and Silverman in their limited survey. Discussing the findings of various studies under 15 different factors, they include that definite prognostic value attaches to the factors of duration of illness, acuteness of onset, presence of obvious exogenic precipitants, confusion and atypical symptoms, "process" symptoms, and subtypes. The survey suggested questionable or no prognostic validity for the factors of age, sex, education, psychosexual history, previous episodes, body type, extroversion-introversion, and heredity. While this review is not exhaustive, it does help to indicate the very limited number of factors on which the research studies are uniform in their findings.

One of the most thorough reviews of the literature on prognosis was published by Blair [1] in 1940. This is particularly good with respect to the European studies, although it is less complete in its coverage of American research. It is primarily concerned with "spontaneous remission" rates as a base line for evaluating efficacy of cardiozol and insulin therapies.

Purpose and Subjects

The purpose of the present investigation was to add to the very limited research data on prognostic factors in schizophrenia. It was a further object of the study to avoid some of the errors and sources of inadequacy of previous studies by using a reasonably large sample of patients, by evaluating their *social adjustment* through direct contact after an interval of some years following initial hospitalization, and by attempting a reasonably exhaustive study of case histories and clinical data in a search for prognostic indicators.

The criterion data on which this study is based were collected as part of an over-all follow-up study of psychiatric patients [5]. In 1949, a graduate social worker trained in research traveled throughout Minnesota to interview patients who had been hospitalized on the psychiatric service of the University of Minnesota Hospitals between 1938-1944. This interval was chosen because it covered the period from the inception of the psychiatric service to the start of extensive utilization of electric and convulsive therapies. Thus, data on the subsequent adjustments of patients evaluated in this study permit appraisal of the effects of psychiatric hospitalization exclusive of the shock therapies which have assumed a major role in recent years.

The social worker succeeded in establishing contact with the patient directly, or with close relatives, friends, and neighbors, or with well-informed agencies in the home community, in 77% (1,261) of the total of 1,638 cases treated during the 1938-1944 interval. In terms of the frequency of the various diagnostic categories, the 77% sample was found to be essentially identical with the total original group of hospitalized patients. Schizophrenics constituted 14.1% of the original group and 16.7% of the group with which follow-up contact was achieved. Of the 16 diagnostic groups represented in the total follow-up sample, the schizophrenics ranked second for percentage of follow-up contact; 87.5% of the 240 patients hospitalized in the interval under study were traced by the social worker. General descriptive data for these patients are reported in Table 2.

A word should be said concerning the nature of the treatment received by the schizophrenic patients. They were hospitalized for an average of 40 days ($M = 39.9$; $\sigma = 30.1$) on a small locked ward (35 beds). The psychiatric service was primarily a diagnostic and teaching unit with a small, eclectically oriented staff. In the assignment of formal diagnoses the official nosological system of the American Psychiatric Association was applied. The usual medical facilities of a psychiatric hospital were available and were used as indicated (e.g., sedatives, hydrotherapy, nursing management of regimen, etc.). Occupational therapy facil-

ities were available and were used extensively in the treatment program; general recreational facilities were extremely limited. Because of the small staff and the generally diagnostic orientation of the unit, psychotherapy was of a most limited nature. No use was made during the interval under study of convulsive or "shock" therapies. It might reasonably be held that recovery and improvement figures for this sample would reflect relatively "spontaneous" remission, and that factors found to differentiate between those with good and those with poor posthospital courses might be used to equate later patient groups for study of the efficacy of specific therapies such as insulin shock.

Procedures

On the basis of interviews and hospital records for those cases in which rehospitalization had occurred, the social worker assigned a rating to each patient to indicate the degree of difficulty he had experienced in his over-all adjustment subsequent to his hospitalization at the university. A five-step rating scale was used, with each step defined as follows:

A. Patient had no further trouble of the type for which he was hospitalized originally.

B. Patients were largely making a good social adjustment. To a small extent they had been bothered by the complaints they presented at the time of hospitalization, but these occurred very infrequently, bothered the patient or his family to a minimal degree, or had gradually diminished and disappeared within the year following hospitalization. These patients may have spent a short time in a state hospital; usually it was immediately following initial hospitalization at the university, and the patient was soon paroled and never returned.

C. Patients were in and out of state hospitals, or worked and made a good adjustment for a time but had relapses of the illness for which they were hospitalized. Although these patients had been well for relatively long periods of time, they usually had two or three episodes of illness.

D. Patients had spent more than half of the time between university hospitalization and follow-up in an institution, incapacitated because of the illness, or functioning on the inadequate level which characterized them at the time of university hospitalization. For the most part, the patient or the patient's family considered him to have a large amount of trouble.

E. Patients had continuous trouble of the type for which they were hospitalized. Many of them were sent from the university hospitals directly to a state institution, and they remained there. Others were cared for at home by their families, but none of them recovered even for a short while.

Table 1 presents the distributions of the follow-up ratings for the total sample of 210 schizophrenics contacted. In comparing the figures of this table with those in the existing literature, it should be noted that the criterion of evaluation applied in this study was not one of degree of psychiatric symptomatology, but rather one of *social adjustment,* and that the interval of posthospital adjustment ranged from five to eleven years.

Schizophrenic patients who fell into the four extreme categories of rated posthospital adjustment (A and B, D and E) constituted the criterion groups in a search for prognostic factors. Adequate case history data and clinical notes were available for all of the A and B cases and for all but seven of the D and E cases. The criterion groups were thus composed of 58 patients who had no further difficulty of the type for which they were originally hospitalized or who were making a good social adjustment (Good group), and 119 schizophrenics who had spent over 50% of the time since their university hospitalization in a mental institution or who had experienced continuous difficulty necessitating their being cared for by their families (Poor group).

The hospital records and case histories of the 178 schizophrenics were next analyzed, with reference to a very detailed schedule for notation of presence or absence, or a rating of the degree, of some 250 items covering personal history, medical history, and development and features of the psychosis.[2] A random selection of records was independently rated on the case history variables by two raters. Complete agreement as to rating was found to obtain on between 80% and 95% of the items in indi-

[2]Copies of the 12-page rating schedule for case history analysis and of the 4-page record form used with it have been deposited with the American Documentation Institute. Order Document No. 4209 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. S., remitting in advance $1.25 for microfilm or $1.25 for 6 x 8 in. photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

This schedule was originally developed by W. Grant Dahlstrom in connection with an unpublished doctor's dissertation [4].

Table 1

Rated Posthospital Adjustment of Schizophrenics Who Received No Specific Form of Therapy

Adjustment rating	*N*	% of group
A (best)	14	7
B	44	21
C	26	12
D	29	14
E (worst)	97	46
Total	210	100

vidual records. Free use was made of an "unknown" category in order not to introduce invalidity by forcing ratings on variables for which data were frequently inadequate. Ratings of the case histories were made without knowledge of which records belonged to which outcome group.

The case histories provided adequate data for a sufficient number of patients to permit statistical analyses of the differential appearance of 200 items in the two criterion groups. Basic data derived from the four parts of the case history analyses are reported in Tables 2-5. While it is clear from inspection that much of the data reported in these tables has no prognostic significance, it is presented in detail as normative statistics which describe the manner in which the clinical material of schizophrenia presents itself. It is hoped that these data may serve as an addition to the textbook descriptions of this psychosis, which generally are unsupported by such surveys.

Of the 200 items statistically tested, only 17 met the 5% level of reliability. While this number of positive findings, and their content, is such as to argue against mere chance [2], the fact that only 17 distinctions should be yielded by 200 items attests to the subtlety of the social and personal factors determining postpsychotic adjustment.

Results

Table 2 presents general descriptive data for the groups with good and with poor posthospital adjustment, and for the two groups combined. The age variable is age at the time of hospitalization. The Good group has both a lower mean age at hospitalization and a reliably more restricted age range than the Poor group ($F = 1.553$; $p < .01$). As is to be expected, these differences are paralleled by the data on "age at onset" (Table 3). The Good group is younger at age of onset and has a reliably smaller variance on this variable than the Poor group ($F = 2.293$; $p < .01$).

As further shown in Table 2, the two outcome groups are *not* differentiated by factors of sex, urban versus rural home, education, religion, marital status, number of children, and number of siblings. While the Good group is characterized by nearly twice the proportion of catatonics as the Poor group, the distribution of the various subtypes in the two outcome groups is not reliably different ($\chi^2 = 9.12$; $p < .10$).

Of the thirty items of personal history recorded in Table 3, only four discriminate between the two groups. The Poor group is characterized by a reliably higher frequency of *excellent* deportment in school ($\chi^2 = 7.82$; $p < .02$). The Good group reveals a reliably higher frequency of church attendance ($\chi^2 = 9.67$; $p < .01$). The marital adjustment of the Good group is clearly superior to that of the Poor group, showing no incidence of "extreme frustration" or "continual conflict," and nearly three times as many "pleasant" marital relationships are recorded for the Good group as for the Poor. Affection toward the mate was manifested by nearly all of the married schizophrenics with good outcome (94.1%), and by only 50% of those in the Poor group ($t = 4.17$; $p < .01$).

In view of some of the prominent thinking concerning psychogenic etiology, it is of considerable interest to note from Table 3 that the two groups are not different with respect to interparental relationships, relationship with either parent, relationship to sibs, or sibling rank (Items 1-5). It is also interesting to note that they are not differentiated by the extent of social participation (Item 13), by their heterosexual adjustment (Item 19), or by their initiative, social intelligence, and poise (Items 22, *e*, *f*, and *g*).

Table 4 reviews the data descriptive of the presenting illnesses of the schizophrenic patients. As noted above, the patients with good outcome were characterized by a reliably smaller range of age at onset (Item 7). The dura-

Table 2

Gross Descriptive Data for Schizophrenics with Good and Poor Posthospital Courses *

Datum	Good group $N = 58$	Poor group $N = 119$	Total group $N = 177$
1. Age	$M = 25.2$	$M = 30.5$	$M = 28.7$
	$SD = 7.0$	$SD = 10.0$	$SD = 9.6$
2. Sex			
Male	47.4%	42.8%	44.4%
Female	52.6	57.2	55.6
3. Home			
Urban	41.3%	55.4%	50.8%
Rural	58.7	44.6	49.2
4. Education	($N = 57$)	($N = 110$)	($N = 167$)
Less than 8 gr.	12.3%	10.0%	10.8%
8 gr. & more	87.7	90.0	89.2
5. Religion			
Lutheran	39.1%	33.7%	35.7%
Other Prot.	23.9	34.9	31.0
Catholic	34.8	22.8	27.1
Jewish	2.2	8.6	6.2
6. Number of Sibs	$M = 4.4$	$M = 4.1$	$M = 4.2$
	$SD = 3.2$	$SD = 3.4$	$SD = 3.4$
7. Marital status			
Married	27.1%	26.8%	26.9%
Single	71.2	68.0	69.1
Divorced	0.0	2.5	1.7
Widowed	1.7	1.6	1.7
Separated	0.0	1.1	0.6
8. Number of children	($N = 17$)	($N = 39$)	($N = 56$)
	$M = 2.2$	$M = 3.1$	$M = 2.8$
	$SD = 2.2$	$SD = 2.4$	$SD = 2.4$
9. Type of schizophrenia			
Catatonic	33.8%	18.4%	23.6%
Simple	18.6	21.8	20.8
Paranoid	16.9	35.2	29.2
Hebephrenic	10.1	9.2	9.6
Unclassified or mixed	20.6	15.4	16.8

* The N's in this and subsequent tables vary with different item comparisons in terms of the number of patients for whom there was adequate information. Where N's are not indicated for the item, data are for the total groups, i.e., $N = 58$, $N = 119$, respectively.

tion of the illness prior to hospitalization was considerably shorter in the Good group than in the Poor group (Item 8, $\chi^2 = 22.2$; $p < .01$). The rate of a history of previous episodes was twice as great in the Poor group as in the Good group (Item 9, $t = 2.83$; $p < .01$). The distribution of the "poor" and "fair" categories of clinical prognosis at time of discharge yields a small but reliable separation of the two outcome groups (Item 11, $\chi^2 = 4.36$; $p < .05$). The Good group had a reliably smaller variation in length of hospitalization than the Poor group (Item 12 *a*, $F = 1.53$; $p < .05$). Finally, the adjustment to hospitalization of the Good group was inferior to that of the Poor group (Item 12 *b*, $\chi^2 = 9.58$; $p < .01$). It is of particular interest to note from Item 12*c* of Table 4 that the long-term adjustment of the schizophrenics was not reliably reflected in their condition at time of discharge; nearly two-thirds of the patients who later made a good adjustment were clinically judged as unimproved when discharged, while 20% of the patients with subsequently poor social adjust-

Table 3

Personal History Data for Schizophrenics with Good and Poor Posthospital Courses

Item	Good group N = 58	Poor group N = 119	Total group N = 177
1. Interparental relat.	(N = 35)	(N = 66)	(N = 101)
Affection	82.8%	72.7%	76.2%
Ambivalence	5.7	6.0	5.9
Indifference	0.0	3.0	1.9
Hostility	11.4	18.2	15.8
2. Paternal relat.	(N = 41)	(N = 86)	(N = 127)
Affection	70.7%	62.8%	65.3%
Ambivalence	9.8	3.4	5.5
Indifference	0.0	5.8	3.1
Rejection	9.8	12.8	11.8
Overprotection	0.0	2.3	1.5
Domination	9.8	12.8	11.8
Jealousy	0.0	0.0	0.0
Dependency	0.0	0.0	0.0
Neglect	0.0	0.0	0.0
3. Maternal relat.	(N = 47)	(N = 81)	(N = 128)
Affection	61.7%	66.6%	64.8%
Ambivalence	2.1	2.4	2.3
Indifference	4.2	0.0	1.5
Rejection	2.1	8.6	6.2
Overprotection	12.8	13.6	13.2
Domination	14.8	8.6	10.9
Jealousy	0.0	0.0	0.0
Dependency	0.0	0.0	0.0
Neglect	2.1	0.0	0.7
4. Sibling relat.	(N = 43)	(N = 83)	(N = 126)
None	4.6%	3.6%	3.9%
Affection	76.7	75.9	76.1
Indifference	4.6	4.8	4.7
Rivalry	11.6	14.4	13.4
Domination	0.0	1.2	0.7
Submission	2.3	0.0	0.7
5. Sibling rank	(N = 34)	(N = 63)	(N = 97)
0	5.8%	4.8%	5.1%
1	26.4	28.6	27.8
2	20.6	19.0	19.5
3	17.6	14.2	15.4
4	14.7	17.4	16.5
5	5.8	6.3	6.1
6	5.8	3.2	4.1
7	0.0	1.6	1.0
8	2.9	0.0	1.0
9	0.0	0.0	0.0
10	0.0	3.2	2.1
11	0.0	1.6	1.0
6. Home			
Poverty	5.0%	10.9%	9.0%
Alcoholism	6.8	6.7	6.8
Invalid	1.6	0.0	0.6
Divorce	1.6	1.6	1.6
Separation	3.4	0.0	1.1
Death of parent	10.2	10.9	10.7
7. School acceptance	(N = 53)	(N = 106)	(N = 159)
Marked hostility	5.8%	2.9%	3.8%
Mild dislike	9.4	12.2	11.3
Indifference	24.5	26.4	25.8
Agreeable	52.8	55.6	54.7
Keen enjoyment	7.5	2.9	4.4
8. School achievement	(N = 56)	(N = 107)	(N = 163)
Repeated failure	7.3%	3.7%	4.9%
Work difficult	14.2	22.4	19.6
Av. performance	58.9	45.8	50.2
Easily earned grades	10.7	24.4	19.6
Accelerated	8.9	3.7	4.4
9. School deportment	(N = 37)	(N = 77)	(N = 114)
Poor record	13.6%	5.2%	7.9%
Usual no. of escapades	21.6	5.2	10.6
Excellent record	64.8	89.6	81.6
10. Occupation			
None	22.0%	17.6%	21.1%
11. Success	(N = 40)	(N = 87)	(N = 127)
Poor	7.5%	18.4%	14.9%
Average	30.0	32.2	31.5
Good	62.5	49.4	53.5
12. Satisfaction	(N = 40)	(N = 86)	(N = 126)
Dislike	12.5%	8.1%	9.5%
Indifference	32.5	48.8	43.6
Enjoyment	55.0	43.0	46.8

Table 3 (Continued)

13. Social	(N = 31)	(N = 54)	(N = 85)
Withdrawal	58.1%	63.0%	61.2%
Ambivalence	12.9	14.8	14.1
Membership	29.0	22.2	24.7
14. Dating	(N = 45)	(N = 100)	(N = 145)
None or little	60.0%	56.0%	57.2%
Average	40.0	44.0	42.8
Very popular	0.0	0.0	0.0
Steady dating	35.6	36.1	35.9
15. Recreation	(N = 55)	(N = 108)	(N = 163)
Predominantly solitary	43.6%	38.8%	40.5%
Mixed	47.2	54.6	52.2
Social	9.2	6.4	7.3
16. Church attendance	(N = 49)	(N = 98)	(N = 147)
Very infrequent	16.3%	41.8%	33.3%
Occasionally	42.8	26.5	31.2
Steady	40.8	31.7	35.3
17. Delinquency	3.4%	2.5%	2.8%
18. Criminal record	5.0%	5.0%	5.0%
19. Heterosexual adjust.	(N = 58)	(N = 113)	(N = 171)
Poor	22.4%	22.1%	22.2%
Fair	31.0	34.5	33.3
Good	46.6	43.4	44.4
20. Adequacy of outlet	(N = 58)	(N = 112)	(N = 170)
Poor	17.2%	28.6%	24.7%
Fair	32.8	29.4	30.5
Good	50.0	42.0	44.7
21. Marital			
None	71.2%	71.4%	71.3%
a. Adjustment	(N = 16)	(N = 34)	(N = 50)
Extreme frustration	0.0%	23.5%	16.0%
Continual conflict	0.0	20.6	14.0
Compatibility	50.0	38.2	42.0
Pleasant	43.8	14.7	24.0
Chief pleasure	6.2	2.9	4.0
	(N = 17)	(N = 34)	(N = 51)
b. Affection to mate	94.1%	50.0%	64.7%
22. Personality			
a. Breadth of interest	(N = 58)	(N = 114)	(N = 172)
Narrow	27.1%	39.4%	35.2%
Some interest outside self	64.4	57.8	60.2
Broad interests	8.5	2.8	4.6
b. Level of aspiration	(N = 58)	(N = 112)	(N = 170)
Limited goals	34.4%	34.8%	34.7%
Interested in improving	56.8	54.4	55.3
High ambition	8.8	10.8	10.0
c. Life plan	(N = 54)	(N = 110)	(N = 164)
Vague	44.4%	47.2%	46.4%
Confused and unrealistic	30.0	27.2	28.0
Fairly clear	25.6	25.6	25.6
d. Stability	(N = 58)	(N = 114)	(N = 172)
Constant fluct.	16.9%	12.2%	13.9%
Moderate variability	42.3	42.9	42.7
Stolid, even tempered	40.6	44.7	43.4
e. Initiative	(N = 58)	(N = 111)	(N = 169)
Apathetic	13.8%	16.9%	16.0%
Appropriate	62.6	48.6	53.2
Energetic	23.6	39.5	30.8
f. Social intelligence	(N = 58)	(N = 109)	(N = 167)
Inept and clumsy	3.4%	11.9%	8.9%
Moderate skill	93.2	86.2	88.8
Adept	3.4	1.9	2.3
g. Poise	(N = 58)	(N = 112)	(N = 170)
Retiring and sensitive	37.2%	38.4%	37.4%
Fairly articulate	50.8	48.2	49.7
Confident in most situations	11.8	13.4	12.8
h. Religiousness	(N = 43)	(N = 78)	(N = 121)
Largely intellectualized and ritualistic	0.0%	6.4%	4.1%
Occasional solace	58.2	53.8	55.4
Dominant source of balance and peace of mind	41.8	39.7	40.5

Table 4

Elements in the Presenting Illness of Schizophrenics with Good and Poor Posthospital Courses

Item	Good group $N = 58$		Poor group $N = 119$		Total group $N = 177$	
	% Patients	% Problems	% Patients	% Problems	% Patients	% Problems
1. Areas of manifestation of illness						
Home	35.2	18.2	45.6	19.6	37.1	19.2
Social	20.4	10.6	26.2	11.2	21.4	11.1
Occupational	25.9	13.4	35.9	15.4	28.7	14.8
Religious	3.7	1.9	6.8	2.9	5.1	2.6
Sexual	16.6	8.6	31.0	13.3	23.0	11.9
Personality	90.6	47.1	87.3	37.4	87.1	40.4
2. General health	($N = 54$)		($N = 108$)		($N = 162$)	
Invalid	0.0%		1.9%		1.2%	
Many minor illnesses	13.0		13.0		13.0	
Usual diseases	18.5		21.2		20.4	
Occasional ailment	40.7		28.7		32.7	
Very good health	27.8		35.2		32.7	
3. Handicaps	6.7%		2.5%		3.9%	
4. Deformities	1.6%		1.6%		1.6%	
5. Weight	($N = 45$)		($N = 79$)		($N = 124$)	
Underweight	2.2%		5.1%		4.1%	
Thin	11.1		19.0		16.1	
Well-proportioned	82.2		67.1		72.5	
Buxom	0.0		6.3		4.1	
Overweight	4.4		2.5		3.2	
6. System involvement						
Cardiorespiratory	3.3%		0.8%		1.7%	
Gastrointestinal	1.6		2.5		2.2	
Genitourinary	1.6		5.8		4.5	
7. Age at onset	($N = 58$)		($N = 114$)		($N = 172$)	
	$M = 24.6$		$M = 26.4$		$M = 25.9$	
	$SD = 7.0$		$SD = 10.6$		$SD = 8.7$	
8. Development of symptoms	($N = 58$)		($N = 116$)		($N = 174$)	
Over a year	25.8%		62.0%		50.0%	
A month to a year	55.2		30.2		38.5	
Less than a month	19.0		7.8		11.4	
9. Previous episodes	18.6%		37.8%		31.4%	
10. Etiology	($N = 55$)		($N = 116$)		($N = 171$)	
Personality	98.1%		99.1%		98.8%	
Family	9.0		11.2		10.5	
Social	10.9		5.2		7.0	
Occupational	10.9		2.6		5.2	
Capacity	3.6		4.3		4.0	
Health	5.4		6.8		6.4	
11. Prognosis	($N = 51$)		($N = 110$)		($N = 161$)	
Poor	43.1%		60.9%		55.2%	
Fair	52.9		36.3		41.6	
Good	3.9		2.7		3.2	
12. Course in hospital						
a. Length	($N = 57$)		($N = 119$)		($N = 176$)	
	$M = 38.6$ days		$M = 40.4$ days		$M = 39.9$ days	
	$SD = 27.6$		$SD = 34.4$		$SD = 30.1$	
b. Adjustment	($N = 53$)		($N = 108$)		($N = 161$)	
Poor	28.2%		13.0%		18.1%	
Fair	49.0		42.6		44.6	
Good	22.8		44.4		37.3	
c. Progress	($N = 56$)		($N = 118$)		($N = 174$)	
No change	62.5%		66.9%		65.5%	
Small improv.	10.7		16.9		14.9	
Decided improv.	26.8		20.3		22.4	

Table 5

Mental Status of Schizophrenics with Good and Poor Posthospital Courses

Item	Good group N = 58	Poor group N = 119	Total group N = 177
1. Delusion	(N = 22)	(N = 44)	(N = 66)
Mild transitory	36.4%	31.8%	33.3%
More extensive	50.0	45.4	47.0
Complex, well-established	13.6	22.7	19.7
Grandeur	3.4%	0.8%	1.7%
Influence	8.4	9.3	9.0
Nihilism	5.0	1.6	2.8
Persecution	35.6	36.1	35.9
Reference	15.2	27.7	23.6
Religion	13.6	14.2	14.0
Self-change	1.6	3.4	2.8
Sex	13.6	5.8	8.5
Somatic	32.2	26.8	28.6
Unreality	5.0	1.6	2.8
2. Hallucination			
Auditory	35.6%	31.0%	32.6%
Visual	32.2	22.6	25.8
Olfactory	1.7	2.6	2.2
Haptic	0.0	0.8	0.6
Gustatory	0.0	0.0	0.0
3. Complaints—hypochondriacal	3.4%	7.6%	6.2%
4. Feelings			
Apathy	15.2%	17.6%	16.9%
Depersonalization	3.4	6.7	5.6
Doubt	0.0	0.8	0.6
Guilt	18.6	16.0	16.9
Inferiority	1.6	4.2	3.4
Isolation	1.6	0.0	0.6
Perplexity	10.2	3.4	5.6
Unreality	6.8	7.6	7.3
5. Orientation			
a. Time	(N = 51)	(N = 110)	(N = 161)
Gross disorientation	19.6%	6.5%	10.6%
Minor errors	9.8	7.2	8.1
Answers well	70.6	86.3	81.3
b. Place	(N = 51)	(N = 111)	(N = 162)
Gross disorientation	9.8%	3.6%	5.6%
Minor errors	9.8	2.8	4.9
Answers well	80.4	93.6	89.5
c. Person	(N = 50)	(N = 111)	(N = 161)
Gross disorientation	8.0%	4.5%	5.6%
Minor errors	4.0	3.6	3.7
Answers well	88.0	91.9	90.7
6. Memory			
a. Past	(N = 46)	(N = 107)	(N = 153)
Vague	6.5%	4.7%	5.2%
Minor errors	6.5	9.3	8.5
Accurate	87.0	86.0	86.3
b. Recent events	(N = 46)	(N = 101)	(N = 147)
Vague	15.3%	9.0%	10.9%
Minor errors	21.7	12.8	15.6
Accurate	63.0	78.2	72.5
c. Recent verbal	(N = 27)	(N = 44)	(N = 71)
Vague	25.9%	22.7%	24.0%
Minor errors	18.5	9.1	12.5
Accurate	55.6	68.2	63.5
d. Recent numerical	(N = 36)	(N = 69)	(N = 105)
Vague	16.6%	11.7%	13.3%
Minor errors	27.8	27.5	27.6
Accurate	55.6	60.8	59.0
7. Retrospective falsification	(N = 9)	(N = 21)	(N = 30)
	44.0%	28.5%	33.3%
8. Thinking capacity	(N = 42)	(N = 98)	(N = 140)
Low	9.5%	9.2%	9.3%
Below average	35.7	36.7	36.4
Average	47.6	41.8	43.5
Above average	7.2	9.2	8.6
High	0.0	3.1	2.1
9. General information	(N = 41)	(N = 93)	(N = 134)
Low	14.6%	5.4%	8.2%
Below average	14.6	34.4	28.3
Average	51.2	43.0	45.6
Above average	14.6	14.0	14.2
High	5.0	3.2	3.7

Table 5 (Continued)

10. Emotional tone	(N = 57)	(N = 114)	(N = 171)
a. Stability			
wide lability	14.0%	8.8%	10.5%
change	14.0	24.6	21.0
stable	72.0	66.6	68.4
b. Depression-Elation	(N = 54)	(N = 112)	(N = 166)
Depressed	55.5%	28.5%	37.3%
Composed	31.5	60.6	51.2
Happy	13.0	10.9	11.9
c. Anxiety	23.7	16.0	18.5
d. Moodiness	3.4	3.4	3.4
e. Irritability	10.2	5.0	6.8
11. Insight	(N = 57)	(N = 117)	(N = 174)
Poor	57.8%	69.2%	65.5%
Fair	40.4	29.9	33.3
Good	1.8	0.9	1.1
12. Stream of thought			
a. Refusal-Flight	(N = 58)	(N = 115)	(N = 173)
Refusal to speak	12.0%	6.0%	8.1%
Retardation	24.1	21.7	22.5
Quiescent	24.1	20.8	22.0
Normal	12.0	27.0	22.0
Verbose	7.0	5.3	5.8
Voluble	7.0	9.6	8.6
Flight of ideas	13.8	9.6	11.0
b. Consist.-Distract.	(N = 51)	(N = 108)	(N = 159)
Consistent	66.6%	61.1%	62.9%
Circumstantial	9.9	10.2	10.0
Rambling	23.5	28.7	27.1
c. Coher.-Incoher.	(N = 53)	(N = 107)	(N = 160)
Coherent statements	90.6%	89.7%	90.0%
d. Impoverishment	6.8	4.2	5.1
e. Echolalia	1.6	0.0	0.6
13. Attitude, manner, behavior			
a. Cooperativeness	(N = 57)	(N = 116)	(N = 173)
Poor	26.3%	14.6%	18.5%
Fair	22.8	32.8	29.5
Good	50.8	52.6	52.0
b. Rapport	(N = 54)	(N = 116)	(N = 170)
Superficial	42.6%	33.6%	36.5%
Friendly	42.6	40.5	41.2
Close	15.8	25.9	22.3
c. Psychomotor			
Stupor	6.8%	2.5%	4.9%
Apathy	50.8	50.4	50.5
Suggestibility	1.6	0.8	1.1
Mimicry	0.0	0.8	0.6
Facetiousness	3.4	0.8	1.7
Querulousness	0.0	2.5	1.7
Hostility	16.9	5.8	9.6
Assaultiveness	11.8	8.4	9.6
d. Neatness	(N = 49)	(N = 103)	(N = 152)
Dishevelled	30.6%	26.2%	27.6%
Neglected	24.6	29.1	27.6
Well-groomed	44.8	44.6	44.7
e. Maturity	(N = 49)	(N = 45)	(N = 94)
Immature	6.9%	0.0%	2.7%
Minor anachronism	0.0	2.2	1.3
Behavior mature	93.1	97.8	96.0
f. Miscellaneous			
Panic	5.0%	2.5%	3.4%
Tension	22.0	8.4	12.9
Tremor	5.0	3.4	3.9
Blushing	1.6	0.0	0.6
Tics	1.6	1.6	1.6
Mannerisms	11.8	18.4	16.3
Hyperkinesis	13.6	11.8	12.4
Catalepsy	10.2	7.6	8.5
14. Symptoms			
a. Affect. cont.	(N = 42)	(N = 93)	(N = 135)
Disharmony			
Appropriate display of emot.	50.0%	43.0%	45.1%
Mild dislocation	40.4	40.8	40.7
Gross splitting	9.6	16.2	14.0
b. Phobias	1.6	4.2	3.3
c. Insomnia	8.4	7.6	7.9
d. Hypersensitive	1.6	1.6	1.6
e. Rigidity	6.8	2.5	3.9
f. Echopraxia	3.4	0.0	1.1

Table 5 (Continued)

g. Stereotypy	15.2	0.8	5.6
h. Litigious	3.4	2.5	2.8
i. Jealousy	1.6	0.8	1.1
j. Fatigue	1.6	2.5	2.2
k. Dreamy states	5.1	1.6	2.8
l. Night terrors	5.1	0.0	1.9
m. Compulsions	3.4	0.8	1.9
n. Egocentric	3.4	5.0	4.5
o. Temper tantrum	3.4	4.2	3.9
p. Amorality	0.0	0.8	0.6
q. Regression	5.1	5.0	5.1
r. Exhibition	3.4	2.5	2.8
s. Indecision	3.4	1.6	2.2
t. Autism	0.0	0.0	0.0
15. Other behavior			
Inappropriate	18.6%	16.0%	16.9%
Silly	11.9	7.6	9.0
Suicidal	5.0	4.2	4.5
Stubborn	23.7	25.2	24.7
Illogical	3.4	2.5	2.8
Suspicious	16.9	23.5	21.3
Impulsive	3.4	5.8	5.0
Childish	5.0	5.0	5.0
Affected	0.0	0.8	0.6
Remorseful	1.6	1.6	1.6
Self-accusatory	5.0	3.4	3.9
Tearful	20.3	6.7	11.3
Nail biting	3.4	1.6	2.2
Insecure	5.0	1.6	2.8
Self-deprec.	3.4	6.7	5.6

ment were considered to have made marked psychiatric improvement.

Over 90 different items pertaining to the mental status and symptomatology of the schizophrenic patients are reported in Table 5. Only six of these items show reliably different frequencies in the two outcome groups. Ideas of reference (Item 6) occurred in a reliably larger proportion of the Poor group ($t = 2.00$; $p < .05$). The Good group showed nearly three times the incidence of gross time disorientation (Item 15*a*) as was found in the Poor group ($\chi^2 = 7.68$; $p < .05$). With respect to emotional tone (Item 20*b*), the Good group was much more frequently described as dull or depressed and much less frequently described as composed when compared with the Poor group ($\chi^2 = 13.27$; $p < .01$). Tension (Item 23*f*) was reliably more frequent in the Good group ($t = 2.30$; $p < .05$.). Stereotypy (Item 24*g*) was practically absent in the Poor group, but noted in 15% of the patients with good post-hospital adjustments ($t = 3.054$; $p < .01$). Finally, tearfulness (Item 25*b*) was over three times as frequent among those schizophrenics with good adjustments as it was in those with poor ($t = 2.40$; $p < .05$).

Some attention should be paid to the failure of all but this very limited number of mental status items to distinguish the two groups. In particular, note is made of the fact that the degree of delusional involvement was not differentially distributed in the two groups, that feelings of apathy had essentially equal incidence, and that appropriate display of affect characterized nearly half of each subgroup. It is equally notable that persecutory, religious, and somatic delusions; degree of insight; and inappropriate, stubborn, and suspicious behavior do not distinguish the two groups from one another. Of much interest, also, is the very low incidence in both groups of symptoms which are commonly considered to be classically descriptive of schizophrenia. For example, in the total sample of 177 schizophrenics, there was a less than 5% occurrence of nihilistic delusions and delusions of self-change and unreality; olfactory, haptic, and gustatory hallucinations; echolalia; stupor, suggestibility, and mimicry; and temper tantrums, exhibitionism, and autism.

A prognostic scale was devised consisting of eight items selected as the most discriminating of the seventeen items reviewed above. The eight items, each of which was given an unweighted unit score in a direction predictive of poor prognosis, were as follows:

1. School deportment: "excellent record; avoided all trouble."
2. Marital adjustment: "source of extreme frustration and emotional turmoil," or "continued mild bickering and conflict."
3. Time taken for symptoms to develop: over a year or more.
4. Presence of previous episodes.
5. Adjustment to hospital: "good; cooperated, showed initiative, carried out assignments well."
6. Presence of ideas of reference.
7. Orientation for time: "answers well within the limits of approximation afforded by sources of information."
8. Emotional tone: "complacent or composed."

Application of this scale to the 177 patients included in the study, using a cutting score of three or more points to identify the cases with poor adjustment, resulted in identification of 59% of the cases in the Good group and 82% of the cases in the Poor group. These figures should be evaluated against the actual 29% occurrence of good outcomes and the 58% incidence of poor outcomes. The degree of improvement over a chance prediction indicated by these figures does not constitute adequate

Table 6

Means and Standard Deviations for the 8-Point Prognostic Scale Applied to Three Different Outcome Groups

Posthospital Adjustment	*N*	*M*	*σ*
Groups A and B	58	2.11	1.24
Group C	26	2.68	1.05
Groups D and E	119	3.68	1.58
Total	203	3.12	1.60

validation of the prognostic scale, however, since the figures represent application of the scale to the same group of subjects on which it was developed. Table 6 gives data on the distribution of the prognosis scale scores in the three outcome groups for subjects whose histories permitted scoring of the eight factors. The statistics of the intermediate group (Group C), which was not used in development of the prognosis scale, satisfy the logical requirements which pertain if there is a reliable, positive relationship between the prognosis scale scores and the criterion of actual posthospital adjustment; i.e., the mean score is intermediate to that of the two extreme groups, and the standard deviation is less than that of either of the two other groups (A and B, D and E) which constitute larger segments of the criterion continuum.

Attention is called to Item 11 of Table 4, which reports the accuracy of clinical prediction of posthospital adjustment. The psychiatric staff appears to have been more pessimistic about this illness in the sample of patients involved than the follow-up data would justify. While nearly one-fourth of the total group had good posthospital courses, the staff predicted only 4% of such outcomes. Furthermore, they predicted poor subsequent histories for over 40% of the patients who were found actually to have achieved good social adjustment. It must be considered that the prognostications of the psychiatric staff may have been more heavily weighted with considerations of degree of continuing psychiatric involvement (symptomatology) rather than with the less technical, but socially crucial, question of adjustment to the community. Nevertheless, the need for factual information as to prognostic factors is clearly illustrated.

Discussion

The findings of this study may be compared with the conclusions reached by Chase and Silverman [3] as a result of a *partial* survey of the literature on prognosis in schizophrenia through 1940. They found short duration of illness, acuteness of onset, presence of confusion in the early symptomatology, and manic-depressive symptoms to be indicative of good prognosis. The findings of the present research are in agreement on these points. Duration of illness in our data is reflected in the relative discrepancies between the ages at onset and ages at hospitalization in the groups with good and with poor outcomes respectively. The onset-hospitalization interval averages .6 years in the Good group and 4.1 years in the Poor group. Acuteness of onset is reflected in the "rate of development" variable (Table 4) and shows slow development of symptomatology to characterize the Poor group. Confusion is indicated in the disorientation for time, which was more common in the Good group, and manic-depressive symptomatology is suggested by the greater occurrence of depression and tearfulness in the Good group.

Chase and Silverman indicate as also prognostically significant, but less differentiating than the above factors, the fact of previous episodes, adequate prepsychotic life adjustment, and progressive hospital course. The findings of this study are directly *opposite* on these points. Thus, occurrence of previous episodes appears as more characteristic of the Poor group; most of the variables pertaining to previous adjustment (*vide* Table 3, Items 10, 11, 12, 13, 14, 15, 16, 17, 18, and 19) do not discriminate the Good from the Poor group, and course in the hospital (Item 12*c*) showed no reliable differences between the two outcome groups.

With respect to failure of sex, intelligence, education, psychosexual history, and age at onset per se to show prognostic significance, this study agrees with the conclusions of the Chase and Silverman [3] survey.

The present research has considered a number of factors not generally covered in previous investigations. There is a set of factors relatively social in nature, or at least reflective of interpersonal relationships, which differen-

tiates the two outcome groups. In brief, the Good group was characterized by poor school deportment, steady church attendance, good marital adjustment, and affection toward mates. General intelligence as evaluated during the mental status examination revealed no marked difference between the two groups, and hence the difference in school deportment appears unrelated to a difference in ability.

What seems at first glance to be a strange inconsistency in the adjustment of the Good group to the two highly organized institutions responsible for the development of social conformity, namely, the school and the church, is possibly accounted for by the relative demand for active participation by the individual as an individual. The seclusiveness and shyness which would handicap the schizophrenic in the schoolroom would be no detriment to his largely spectator role in church. This explanation, of course, presumes typical prepsychotic personality features and is open to argument on that basis. It is further weakened by the fact that in the present study "social withdrawal," as rated on the basis of case histories, revealed no differential frequency in the two groups! The fact of positive association between good marital adjustment with affection toward the mate, and good posthospital course, seems most reasonable.

It may be pointed out that the poorer hospital adjustment of the members of the group with good outcome is probably related to their tendency toward briefer hospitalization, both of these factors suggesting more acute disturbance in the Good group; this would be in line with usual prognostic thinking and with the greater frequency of the catatonic type in the Good group. To some extent it may be assumed that the greater incidence of stereotypy, tension, and time disorientation in the group with good posthospital adjustment is associated with the greater incidence of the catatonic subtype in this group. Similarly, the occurrence of previous episodes and ideas of reference as more characteristic of the Poor group may reflect the slight predominance of the paranoid type in this group. However, it should be recalled that the over-all distribution of the subtypes was not different for the two outcome groups, and the prognostic factors that have been reviewed in detail may not be considered to simply stem from subtype variation.

The validity of the findings of this study rests heavily on the thoroughness and accuracy of the case histories, which are functions of the skill and care with which they were compiled by the medical, psychological, and social work personnel. It is perhaps unfortunate that these were routine and not "research" histories and that they were not recorded with the guidance of the schedule of items subsequently applied to their analysis. However, the essential casualness of the history taking, in addition to the fact that items were rated only when the data were considered adequate, may introduce a type of clinical validity and permit a generalization of our prognostic factors that might not be the case if the original history data were more "scientific." Quite positively, it is believed that the nature and extent of the follow-up study have provided more accurate and socially pertinent criteria of posthospital adjustment than have previously been applied in studies of prognosis.

The paucity of careful follow-up studies involving reasonably large samples, extended postdischarge intervals, and direct evaluation of subsequent adjustments of patients may account for the generally pessimistic orientation which is held with respect to prognosis in schizophrenia. This pessimism is reflected in the clinical prognoses which were assigned to the patients in this study and appears to have been unjustified in view of their subsequent social adjustments. It is hoped that data of the type presented in this paper may contribute to greater exactness in prognosis, not only by indicating the actual frequency of favorable and unfavorable long-term outcomes, but also by suggesting historical and clinical factors which carry prognostic import. Certainly, good information as to basic rates of remission in the absence of special therapeutic efforts is essential to any reasonable test of the efficacy of particular therapies, such as insulin coma. In establishing control and experimental groups for testing therapeutic efficacy, attention must be given to adequate matching on variables, affecting basic prognosis. As an example of this, discrepancies in the remission rates reported for samples of schizophrenics similarly treated

might be explained in some instances by marked dissimilarity of the samples with respect to diagnostic subtype composition, if these data were reported by the investigators.

Summary and Conclusions

In 1949, a graduate social worker with special research training and experience attempted to establish direct contact with all patients hospitalized on the psychiatric service of the University of Minnesota Hospitals between 1938-1944, an interval when no use was made of insulin or electroshock therapy. Of the original sample of 240 schizophrenics, follow-up data were obtained on 210 patients. These were rated on a five-step scale for their over-all posthospital social adjustment. Twenty-eight per cent of this sample were found to have had no further difficulty or were found to be making a good adjustment (Good group), and 60% were found to have spent over half of the post-University interval in a mental institution or had experienced difficulty which necessitated continuous care by their families (Poor group).

The Good and the Poor outcome groups were rated and contrasted with respect to 200 items pertaining to their general history, previous personality, and presenting illness. Only 17 of the 200 items differentiated the two groups at the 5% level of probability. This number of positive findings from such a large series of tests is sufficient to reject the null hypothesis with respect to the frequency of the items in the populations represented by the two samples. However, the smallness of the number of differentiating items is much more impressive than the fact of statistical reliability. To the degree that the posthospital adjustments of schizophrenics are determined by factors of a historical, personal, and social nature, these factors would appear to be of an extremely subtle nature. It would appear that discovery of those factors in the structure of family relationships — the physical aspects of the home environment, and the neighborhood and community attitudes toward the returned patient — that may be important in determining the quality of his adjustment are highly elusive, and their detection would require a much more intensive clinical evaluation of the posthospital course than has as yet been achieved. The other possibility suggested by the findings of this study would be that subsequent social adjustment of the schizophrenic is determined by factors of a largely endogenous nature that are inherent in the disease process.

With respect to prognostic indicators, the patients with good posthospital adjustments were differentiated from those with poor adjustments by the factors of: shorter duration of illness prior to hospitalization; more rapid development of symptomatology; absence of previous episodes; poor school deportment; steady church attendance; good marital adjustment and affection toward the mate; stereotypy; tension; depression; tearfulness; disorientation for time; and absence of ideas of reference.

Received October 2, 1953.

References

1. Blair, D. Prognosis in schizophrenia. *J. ment. Sci.*, 1940, **86**, 378-477.
2. Brozek, J., & Tiede, K. Reliable and questionable significance in a series of statistical tests. *Psychol. Bull.*, 1952, **49**, 339-341.
3. Chase, L. S., & Silverman, S. Prognostic criteria in schizophrenia: a critical survey of the literature. *Amer. J. Psychiat.*, 1941, **98**, 360-368.
4. Dahlstrom, W. G. An exploration of mental status syndromes by factor analytic techniques. Unpublished doctor's dissertation, Univer. of Minnesota, 1949.
5. Hastings, D. W., Hathaway, S. R., & Bell, Dorothy. Follow-up study of psychiatric patients. *Univer. Minn. Hosp. Bull.*, 1951, **23**, 149-159.
6. Malamud, W., & Render, N. Course and prognosis in schizophrenia. *Amer. J. Psychiat.*, 1939, **95**, 1039-1057.
7. Pollock, H. M. Outcome of mental diseases in the United States. *Ment. Hyg., N. Y.*, 1925, **9**, 783-804.
8. Romano, J., & Ebaugh, F. G. Prognosis in schizophrenia; a preliminary report. *Amer. J. Psychiat.*, 1938, **95**, 583-596.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Direction of Aggression as a Prognostic Variable in Mental Illness

Dorothy A. Feldman
University of Pittsburgh

G. R. Pascal, and C. H. Swensen
University of Tennessee

The differential prognosis of paranoid and depressive disorders has a long history in psychiatric literature [2]. However, predictions based on these diagnostic categories have not been very reliable. In some instances, paranoids do improve and depressed patients do not [5, 8]. That the current classificatory system is unreliable is generally conceded [2, pp. 86-87], and the need for variables which are reliable and which have prognostic value for the outcome of treatment is clear. Taking its cue from the classical differentiation between paranoid and depressive disorders [3] and from studies by Albee [1] and Smith [9], the present study considers direction of aggression as a parameter of mental illness and its relation to prognosis. As part of a larger research project [6], this study also attempts to avoid one of the major weaknesses of earlier prognostic studies in that, as far as possible, other variables of known prognostic significance are held constant while direction of aggression is studied.

Method and Procedure

A previous paper described the selection of a sample of 486 patients from the population of hospitalized mental patients at the Western Psychiatric Institute and Clinic [6]. There were 229 males and 257 females in the sample. Diagnostic categories were represented as follows: 264 with schizophrenia, 58 with affective psychoses, 17 with "other psychoses," 95 with psychoneuroses, and 52 with behavior disorders. Status was categorized as improved or unimproved, according to follow-up information one year after discharge. There were 234 improved and 252 unimproved individuals. Every practical effort was made to insure the reliability of the "improved-unimproved" dichotomy, and evidence was presented to indicate the validity of the judgments [6]. The sample contained no subjects younger than 15 or older than 55 years and no known cases of cortical damage or mental deficiency. No patients had a serious somatic illness, and none had a history of recovery from previous alienations.

Guided by the literature and the feasibility of obtaining reasonably reliable measurements, we selected 21 variables and related them, by appropriate statistical techniques, to status one year after discharge from the hospital. Of these 21 variables, the following 9 were found to be significantly related to the defined outcome of treatment: affective expression, orientation, diagnosis, onset, duration of illness, precipitating stress, type of treatment, marital status, and direction of aggression. The intercorrelations of these variables were also investigated. Then an attempt was made to hold eight of the relevant variables constant while investigating the relationship between status and the ninth variable, direction of aggression. In other words, the group of improved and unimproved patients were of the same size, contained the same number of diagnostic categories, and had no statistically significant differences between their means on length of illness, type of onset, stress, and so on. Such matching, which was deemed best for greater precision, reduced the original sample of 486 to a controlled sample of 100 improved and 100 unimproved subjects.

Direction of aggression was measured by a nine-point rating scale, details of which are described elsewhere [4]. It was found to have a reliability coefficient of .83 when two clinical psychologists independently rated 40 case his-

tories selected at random from the hospital files. The nine points along the continuum from extremely extrapunitive to extremely intropunitive behavior are presented below.

Point one: homicidal attack. The use or threatened use of lethal weapons. Attempts to strangle or maim. Blame projected onto others in conjunction with this behavior.

Point two: combative behavior. Assaultive, extremely destructive, while projecting blame. This point may include a previous suicidal attempt which was obviously aggressive in nature, followed by assaultive behavior.

Point three: verbal abusiveness, or systematized paranoid delusions expressed with or without verbal abuse.

Point four: blame projected onto physical complaints, as long as the hypochondriasis is not bizarre or depersonalized. Vague, unsystematized projection. Ambivalent behavior, with occasional aggressive activity toward others.

Point five: ambivalent, impunitive. This point includes obsessive-compulsive behavior and incoherent states.

Point six: mild depression. Ambivalent behavior with some evidence of depression, as in apathy with occasional tears. Bizarre hypochondriacal complaints.

Point seven: depressed. Feelings of guilt, self-blame expressed, with no projection of blame.

Point eight: deeply depressed. Continuous verbal abuse of the self, expressing sin, guilt, unworthiness. This point may include expression of suicidal wishes and suicidal attempts of a milder type than in point nine.

Point nine: serious suicidal attempts and extreme self-condemnation.

Results and Discussion

The ratings on direction of aggression were found to have a moderate positive skewness when the distribution of the total sample was plotted. This skewness, however, was not considered sufficient to prevent the use of normal curve statistics. A significant biserial coefficient of correlation of .67 was obtained between the ratings and status one year after discharge from the hospital (see Table 1). Those patients in the total sample who received low ratings tended to remain unimproved after treatment. When the relationships between other significant variables and direction of aggression were studied, many showed significant intercorrelations. Table 2 summarizes these relationships. Patients within rating scale points one to three tended to have had less precipitating stress, less affective display, a longer period of illness, an insidious type of onset, and to be unmarried, disoriented, and to carry a diagnosis of schizophrenia. The diagnostic categories for those in the high ranges of the scale tended to be manic-depressive and involutional psychoses, and reactive depressions. Those earning middle ratings were most often diagnosed as having affective psychoses or psychoneuroses. The middle and high groups on the scale showed more affect, greater stress, more pronounced onset, very little disorientation, and shorter duration of illness.

Table 1

Summary Statistics for Total Group and Controlled Sample Group

Statistic	Total group	Controlled sample group
N	486	200
Mean†		
Improved	5.02	4.20
Unimproved	3.16	3.69
Standard deviation		
Improved	1.72	1.28
Unimproved	.95	1.33
t	14.31*	2.83*
r_{bis}‡	.67*	.23*

* Significant at .01 level.
† In rating scale points.
‡ Between direction of aggression and outcome.

Table 2

Relationship between Direction of Aggression and Significant Variables ($N = 486$)

Variable	Statistic	p
Affective expression	$r = .24$	.01
Orientation	$\chi^2 = 12.8$	.01
Diagnosis	$\chi^2 = 132.9$	.01
Type of onset	$r = .20$	.01
Duration of illness	$\chi^2 = 24.3$	.01
Precipitating stress	$r = .23$	.01
Treatment	$\chi^2 = 18.9$	.01
Marital status	$\chi^2 = 9.4$	$>.05$

Table 3 presents a summary of the results of matching the improved and unimproved groups on eight of the significant variables, while the aggression variable was left free to vary. The biserial coefficient of correlation was .67 before matching, and .23 when other relevant factors were held constant, as is shown in Table 1. The

Table 3

Differences between Significant Variables in the Controlled Sample Group

Variable	Improved	Unimproved	Statistic	p
Affective expression				
Mean rating	5.38	5.26	$t = .48$	$>.05$
Orientation				
Oriented (N)	91	86	$\chi^2 = 1.23$	$>.05$
Disoriented (N)	9	14		
Type of onset				
Mean rating	5.11	5.31	$t = .77$	$>.05$
Duration				
Mean, in months	32.4	28.6	$t = .96$	$>.05$
Precipitating stress				
Mean rating	3.03	2.90	$t = .52$	$>.05$
Type of treatment				
Non-physical (N)	30	35	$\chi^2 = .57$	$>.05$
Physical (N)	70	65		
Marital Status				
Single (N)	56	55		
Married (N)	41	34	$\chi^2 = 5.23$	$>.05$
Divorced	3	11		
Diagnosis (N)				
Schizophrenia	45	45		
Affective psychoses	11	11		
Other psychoses	5	5	None	
Psychoneuroses	25	25		
Other nonpsychoses	14	14		

means for the improved and unimproved groups were 4.20 (SD, 1.28) and 3.69 (SD, 1.33), respectively, with a t of 2.83. Although this difference is still significant at the 1 per cent level of confidence, it represents quite a loss in predictive power. However, the retention of a significant relationship to outcome of treatment after the rather stringent method of partialling employed here indicates the probable potency of direction of aggression as a prognostic variable.

The method used here appears to serve two functions. The first is concerned with techniques of studying a single, well-defined variable while controlling the concomitant influence of related factors. The second function is that of suggesting the type of objective measurement which may facilitate comparisons of many studies. A single example may illustrate this point. Paster and Holtzman [7] reported a good prognosis for "paranoids and catatonics," results which at first seem to contradict the present findings. However, careful reading reveals essential agreement. Their improved group consisted of "paranoid schizophrenics who were dull and apathetic" or "manic-depressives with paranoid trends" and "catatonic schizophrenics who were not difficult to manage." It is obvious that one cannot fit all paranoid reactions into one paradigm or all depressive behavior into another. The use of a scale for rating direction of behavior may help to reduce the many differences in criteria which Bellak deplores [2].

Summary

A group of 486 patients hospitalized for "nonorganic" mental illnesses were categorized as improved and unimproved one year after discharge from the hospital. Twenty-one variables, extracted from the case history material, were studied for their relationship with follow-up status. Nine pertinent variables were found to be significantly related to outcome of treatment. One of these, direction of aggression, a rated variable, was found to have a significant relationship with follow-up status when the effects of the other significant variables were partialled out. It was concluded that direction of aggression, as defined, was a reliable prognostic variable. The patients who tended to direct blame or hostility exclusively onto the environment appeared to have a poorer prognosis for status one year after discharge from a mental hospital.

Received September 8, 1953.

References

1. Albee, G. W. Patterns of aggression in psychopathology. *J. consult. Psychol.*, 1950, **14**, 465–468.
2. Bellak, L. *Dementia praecox.* New York: Grune & Stratton, 1948.
3. Blair, D. Prognosis in schizophrenia. *J. ment. Sci.*, 1940, **86**, 378–413.

4. Feldman, Dorothy A. Studies of prognostic criteria in case-histories of hospitalized mental patients. II. Direction of aggression. Unpublished doctor's dissertation, Univer. of Pittsburgh, 1952.
5. Kalinowsky, L. B., & Hoch, P. H. *Shock treatments.* New York: Grune & Stratton, 1946.
6. Pascal, G. R., Swensen, C. H., Feldman, Dorothy A., Cole, Mary E., & Bayard, Jean. Prognostic criteria in the case histories of hospitalized mental patients. *J. consult. Psychol.,* 1953, **17,** 163–171.
7. Paster, S., & Holtzman, S. C. A study of one thousand psychotic veterans treated with insulin and electric shock. *Amer. J. Psychiat.,* 1949, **105,** 811–814.
8. Rickles, N. K., & Polan, C. E. Causes of failure in treatment with electroshock. *Arch. Neurol. Psychiat.,* 1948, **59,** 337-346.
9. Smith, Dorothea. An analysis of aggression among patients in a mental hospital. Unpublished doctor's dissertation, Univer. of Pittsburgh, 1951.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Prognostic Significance of Precipitating Stress in Mental Illness[1]

Mary E. Cole[2]
Psychological Service of Pittsburgh

C. H. Swensen,[2] and G. R. Pascal[2]
University of Tennessee

Precipitating factors or precipitating stress is frequently reported in the literature as having prognostic value in mental illness. The consensus is that the presence of some precipitating factor, especially one of a fairly severe nature, makes for a better prognosis in mentally ill patients than does the absence of a precipitating factor. The rationale for using precipitating stress as a prognostic indicator is that persons who succumb only to fairly severe precipitating factors have a greater stamina or resistance to ordinary life stress than do those who become ill with little or no precipitating stress. This greater resistance to severe stress implies that, when such patients are removed from the stressful situation, as in hospitalization, they have a better chance for a remission of symptoms than would be true if they had less resistance to stress.

Many studies have been made of precipitating stress, and a preceding article [6] contains an extensive bibliography. These studies can be summarized very briefly by stating that most of them agree that the presence of precipitating stress is generally indicative of a better prognosis than the absence of precipitating stress. In most studies little attention is given to statistical treatment of the data beyond reporting percentages. In none of the studies which dealt with multiple prognostic variables was any attempt made to study one of the significant variables while controlling the others. These latter two considerations are the basis for the present research.

[1] Based upon a dissertation submitted to the Graduate School of the University of Pittsburgh by the senior author in partial fulfillment of the requirements for the Ph.D. degree.

[2] At the University of Pittsburgh when this research was done.

Procedure

The case records used in this study were those of the Western Psychiatric Institute and Clinic (WPIC), University of Pittsburgh. The records of all patients admitted to this hospital since November 1942 and discharged not later than July 1950 (to allow a minimum of one-year follow-up) were examined. There were 1,118 admissions and discharges between these dates. The following criteria were then applied to these records for selective purposes: age 15 to 55, lack of known cortical damage, absence of mental deficiency, absence of incapacitating somatic illness, and absence of previous attacks of mental illness. The last criterion requires further explanation. Since WPIC is primarily a referring hospital, most of its patients come from other hospitals. Therefore, the criterion "absence of previous mental illness" means that upon admission to WPIC the patient was suffering from one continuous attack of mental illness without known prior attacks which resulted in referral to a psychiatrist. Application of these criteria left 486 cases available for further study.

Status one year after discharge was categorized as "improved" or "unimproved." A patient was considered improved if, one year after discharge from the WPIC, he was working at a level of efficiency roughly equivalent to that which obtained prior to illness. He was considered unimproved if, one year after discharge from the hospital, he was in another mental hospital, or was manifesting other behavior which, when judged by comparison with expectations about a person of like status, was indicative of inability to cope with his environment.

Twenty-one variables were selected for study in the sample of 486 cases. The variables were: sex, age at admission, marital status, religion, nationality, occupation, economic level, body type, education, intelligence, affective expression, orientation, direction of aggression, diagnosis, patient's age at onset of illness, type of onset, duration of illness, total hospitalization, status on discharge, precipitating stress, and treatment. A more complete treatment of the measurement of each of these variables and their relationship to outcome of illness was reported by Pascal *et al.* [6]. However, since precipitating stress is the special province of this paper, it will be dealt with here in detail.

Rating of Precipitating Stress

A nine-point scale was constructed for rating precipitating stress. Point 1 on the scale represented little or no stress, and point 9 represented severe stress. A condensation of the final scale used to rate the cases is given below. The full scale may be found in Cole's dissertation [3].

Point 1. Little or no precipitating stress.

Included here are those cases in which no precipitating factor was found. Many of these will have case histories of a lifetime of stress and poor adjustment with no period of relatively stable adjustment. In other cases, adjustment may have been fairly adequate, with the mental illness precipitated by some trivial, often temporary, kind of stress which the ordinary person frequently encounters and copes with easily.

Point 3. Relatively mild precipitating stress.

Included here are occurrences which are admittedly somewhat stressful, but which are so common to a great many people that they may be regarded as ordinary life stresses. Also, for most people, the results from this amount of stress are usually temporary in nature. General examples are: induction into military service of a young single man, death of a parent when one is an adult, birth of a child, illness in the family, and mild physical illness or minor injury to the patient.

Point 5. Moderate precipitating stress.

Included under this point are such precipitating factors as induction into military service of a man with family responsibilities, heavy financial responsibilities, moderate illness or injury to the patient or his family, poor marital adjustment, worry about a job, chronic difficulties with in-laws, and the like. Also included in Point 5 are combinations of two or more instances of milder stress, usually occurring within a relatively brief period of time so that they have the cumulative effect of moderate precipitating stress.

Point 7. Relatively severe precipitating stress.

This category of the scale includes instances of environmental stress more severe than the average person may ever experience, and usually involves the loss or threatened loss of something very important to the individual's security. It is exemplified by the threatened loss of a loved one through serious illness or injury, rejection by a loved one, or loss of livelihood after a period of good adjustment. It will also frequently include a combination of mild or moderate stresses occurring together and resulting in the equivalent of severe precipitating stress.

Point 9. Extraordinary, extreme precipitating stress.

This category includes precipitating factors so extreme as to be overwhelming, such as loss of a child or children, or of a spouse where the relationship has been a very close and dependent one. It would also include the complete disruption of a well-adjusted life pattern through some catastrophe or disaster striking suddenly, or the same effect caused by a period of severe, mounting environmental stresses following a long period of stable adjustment. In any event, a case history rated point 9 would be the net result of a far greater amount of precipitating stress than the average person would ever experience.

It will be noted that only five points (points 1, 3, 5, 7, and 9) on the scale were defined. The even-numbered points (points 2, 4, 6, and 8) were used as intermediate points for the rating of cases that appeared to fall between the defined points on the scale.

After the scale points were defined and illustrated, forty cases were selected and independently rated by two clinical psychologists for amount of precipitating stress, using the scale just described. Their ratings correlated .90, indicating fairly good reliability.

The relationship between the variables and outcome. Eleven of the variables (affective expression, orientation, direction of aggression, diagnosis, type of onset, duration, hospital stay, marital status, status on discharge, precipitating stress, and treatment) are significantly related to outcome of illness at the .01 level or better. A more complete treatment of each of these variables is contained in a previous paper [6].

Method used to control the significant variables. The crucial part of this study lies in the selection of the controlled sample population, and in the controlling of all variables significantly related to outcome except the variable

of precipitating stress. Precipitating stress was permitted to vary, and appropriate statistics were then used to determine whether this variable alone was related significantly to status one year after discharge.

Since the criterion used throughout was status one year after discharge, categorized as improved or unimproved, the 486 cases were divided into an improved group of 232 and an unimproved group of 254. These two groups were then matched for the variables found to be significantly related to outcome, other than precipitating stress, by eliminating cases until there were no significant differences in means between the improved and unimproved groups with respect to each of these variables. These controls reduced the improved and unimproved groups to 90 cases each. The differences between the final groups on the significant variables are contained in Table 1.

Table 1

Significance of Differences between the Controlled Sample Groups on the Controlled Variables

Variable	Statistic	p value
Affective expression	$t = .14$	$>.05$
Orientation	$\chi^2 = 3.24$	$>.05$
Direction of aggression	$t = .00$	$>.05$
Type of onset	$t = .50$	$>.05$
Duration of illness	$t = .58$	$>.05$
Total hospital stay	$t = .63$	$>.05$
Treatment	$\chi^2 = .00$	$>.05$
Diagnosis	$\chi^2 = .00$	$>.05$
Marital status	$\chi^2 = .10$	$>.05$

As a further check, the final controlled sample groups were compared with respect to the other ten variables not found to meet the level of confidence of .01 decided upon. It was found that the controlled sample groups did not differ significantly in means for any of these nonsignificant variables.

Results and Discussion

When the controlled sample groups were finally matched on the significant variables, they were compared with respect to precipitating stress, which was the only factor left free to vary. It was found that there were no statistically significant differences between the improved and unimproved groups in amounts of precipitating stress. Table 2 shows the differences between the controlled sample groups in precipitating stress. A t of .35 was obtained, which is not statistically significant. These results were strikingly different from those obtained in the original total sample population of 486 cases, where precipitating stress was found to be related to status one year after discharge at the .01 level of confidence. It will be noted that the mean rating for the total improved group was 3.23, and for the unimproved group, 2.51. However, in the controlled sample groups, the mean rating for the improved group was 2.92, and for the unimproved group it was 3.01. The difference between the total groups was significant, but the difference between the controlled groups was not. The average amount of precipitating stress for this particular sample of hospitalized patients fell at point 3 on the scale. Point 3 includes what might best be described as ordinary life stress: those situations which admittedly are somewhat stressful, but which occur so commonly and are withstood by such a

Table 2

Summary Statistics for the Total Group and the Controlled Sample Group

Statistic	Total group	Controlled sample group
N	486	180
Mean*		
Improved	3.23	2.92
Unimproved	2.51	3.01
Standard deviation		
Improved	1.91	1.60
Unimproved	1.60	1.81
t†	4.43§	.35
r_{bis}‡	.25§	−.04

* In rating scale points.
† Between improved and unimproved groups.
‡ Between stress and outcome.
§ Significant at .01 level.

Table 3

Relationship of Precipitating Stress to Other Significant Variables

Variable	Statistic	p value
Affective expression	$r = .18$	.01
Orientation	$\chi^2 = 6.5$	.05
Direction of aggression	$r = .23$	.01
Diagnosis	$\chi^2 = 12.9$	.02
Type of onset	$r = .11$	.05
Duration of illness	$\chi^2 = 19.3$	.01
Treatment	$\chi^2 = .81$	.50
Marital status	$\chi^2 = 22.24$	.01

large proportion of the population that on this scale they were rated as mild.

To show the difference between the total sample and the controlled sample groups with respect to the prognostic value of the variable, precipitating stress, Table 3 is included, showing the relationships between the eight significant variables and precipitating stress. An examination of these relationships reveals that precipitating stress is significantly related, with a probability value of .01, to the following four variables: affective expression, marital status, direction of aggression, and duration of illness. Relationships approaching significance (less than .05 but greater than .01) are found with orientation, diagnosis, and type of onset. It would appear from these results that precipitating stress is not a pure variable, but rather, that it depends for what prognostic value it has on its relationships to other significant variables.

Most of the significance found between amount of affective expression and precipitating stress derived from the upper end of the precipitating stress scale. That is, patients having more than ordinary life stress as a precipitating factor tended to be rated high on affect. Those patients with very little precipitating stress varied in the amount of affective expression.

With regard to marital status, more single persons became mentally ill with very little precipitating stress than did married ones. However, more married people broke down under severe to extreme precipitating stress than did single people. This latter finding might be interpreted in the light of the additional personal and financial responsibilities that go with marriage, thereby creating more opportunities for stressful situations to occur.

The closest relationship between precipitating stress and any of the significant variables was that found with direction of aggression. Where precipitating stress was high, there was a tendency for aggression to be turned toward the self. Where precipitating factors were few, aggression was found to be more often projected outward toward the environment. Direction of aggression was also significantly related to diagnosis, which perhaps, in turn, gives the approach to statistical significance found between precipitating stress and diagnosis.

As might be expected, duration of illness was found to be significantly related to precipitating stress. In those cases of very long duration, little or no precipitating stress was found, and where the illness was of relatively brief duration, greater amounts of precipitating stress were found.

Summary

Using predetermined criteria, 486 case records were selected from the files of the Western Psychiatric Institute and Clinic. These records comprised all cases admitted up to July 1950 which met certain criteria. All were records of patients between the ages of 15 to 55 who were without mental deficiency, without known damage to the cortex, without incapacitating somatic illness, and without previous attacks of mental illness. The patients' status one year after discharge from the hospital was determined and categorized as improved or unimproved. The records were examined and 21 variables extracted which were, by appropriate statistical techniques, related to status one year after discharge. Of the 21 variables, 11 were found to be of prognostic significance at the .01 level of confidence. These were: affective expression, orientation, direction of aggression, diagnosis, type of onset, duration of illness, length of hospital stay, marital status, status on discharge, precipitating stress, and amount of treatment.

This paper is concerned with precipitating stress as a prognostic variable when all other significant variables are controlled A nine-point scale was constructed for rating precipitating stress. Appropriate reliability and validity checks of the scale were made before it was used on the total population of cases.

After all 486 cases in the total sample population were rated on the scale for precipitating stress, they were divided into two criterion groups: those improved one year after discharge from WPIC, and those who remained unimproved one year after discharge. There were 232 in the improved group and 254 in the unimproved. These two groups were then matched with respect to all the variables found to be significantly related to outcome by with-

drawing cases until there were no significant differences between means and standard deviations for each variable between the improved and unimproved groups. These controlled sample groups numbered 90 improved and 90 unimproved cases. A further check on these two groups compared the means of the ten variables which did not meet the criterion of a probability value of .01, and no significant differences were found between the controlled sample groups on any of these variables. Thus, with all significant and nonsignificant variables controlled, the variable of precipitating stress was the only one left free to vary. The result was that no statistically significant differences were found between the improved and unimproved groups with respect to the amounts of precipitating stress involved in their illnesses.

Received August 11, 1953.

References

1. Bellak, L. *Dementia praecox.* New York: Grune & Stratton, 1948.
2. Cheney, C. W., & Clow, H. E. Prognostic factors in insulin shock therapy. *Amer. J. Psychiat.,* 1941, **97**, 1029–1039.
3. Cole, Mary E. Studies of prognostic criteria in the case histories of hospitalized mental patients: III. Precipitating stress. Unpublished doctor's dissertation, Univer. of Pittsburgh, 1952.
4. Hamilton, D. M., & Wall, J. H. The hospital treatment of dementia praecox. *Amer. J. Psychiat.,* 1948, **105**, 346-352.
5. Hunt, R. C. Relation between precipitating situation and outcome in manic-depressive psychosis. *Amer. J. Psychiat.,* 1938, **95**, 65-72.
6. Pascal, G. R., Swensen, C. H., Feldman, Dorothy A., Cole, Mary E., & Bayard, Jean. Prognostic criteria in the case histories of hospitalized mental patients. *J. consult. Psychol.,* 1953, **17**, 163-171.
7. Strecker, E. A., & Willey, G. F. Prognosis in schizophrenia. *J. ment. Sci.,* 1927, **73**, 9-39.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Patient Cancellations Following a Sudden Cancellation by the Therapist

Miriam Crowley McCue

VA Mental Hygiene Clinic, Lowell, Massachusetts

Psychotherapists and case workers often express the opinion that their canceling of a scheduled appointment will be followed by the patient's cancellation of the next appointment. It is not unusual to hear that a patient canceled, or failed to keep, an appointment "in retaliation" following a cancellation by the therapist. Most therapists will agree that the meaning of the therapist's cancellation depends on the patient and upon his relationship with the therapist at the time. Some patients may be anxious to see the therapist after his absence, and their reasons for this are varied. Others may consciously or unconsciously feel that "if he can cancel, I can too." Still others may think that the therapist is not interested in them, and they may stay away with hurt feelings, or stay away because their own interest begins to lag after an interval without treatment.

Resistance to treatment may follow the therapist's absence, and it may appear in various forms of behavior. One of the most effective modes of resistance is canceling the appointment, for this precludes any interaction between the therapist and patient. If a patient resents his therapist's absence, yet keeps his appointments regularly and on time, the therapist may note resistance in the manner of the patient or in the content of the interview material. The more the patient conceals his resentment, the more difficult is the task of recognizing overt and covert signs of hostility.

Since many varieties of response are possible for the patient, several avenues are open for the study of what sudden cancellations mean to him. It seems in order to start with a simple objective approach and determine first just how often patients cancel appointments after their therapists have canceled appointments with them.

If the therapist is uneasy about canceling, such a count will tell how valid is his concern. Perhaps therapists should ask themselves why they believe that patients will cancel after therapists do. If the likelihood of a patient's acting-out is great, the therapist's anxiety may increase — which only complicates treatment.

The present study was undertaken to discover whether or not patients actually do cancel appointments after their therapist cancels, as many therapists seem to believe.

Method

Subjects. The data were obtained from the attendance cards of 94 male veterans of World War II who were currently or previously in treatment at a mental hygiene clinic in a large mill city in New England. There was little variation in the age, education, and occupational level of the patients. Most of them had weekly appointments, although some were seen twice a week, and some once or twice a month. Therapists of these patients were seven staff members of the clinic, specifically two psychiatrists, four social workers, and one psychologist.

Procedure. The appointment records were examined for instances in which a therapist canceled an appointment with a patient. Only unexpected cancellations were considered, i.e., those of which patients did not know until the day of the appointment. The first such cancellation was the only one counted for any one patient. The experimental group, *E,* made up of 47 *S*s, was drawn from these cases.

A control group, *C,* consisted of 47 *S*s who *kept* the appointment which matched the ther-

apist-canceled appointment of one of the *E* group members. Thus 47 pairs of *S*s were matched for the number of the appointment from the start of therapy; this shall be referred to as the "matched appointment." Matched appointments ranged from the fifth one after therapy started to the seventy-eighth, the median being the sixteenth. The two groups were further matched patient for patient, each pair having the same therapist, approximately the same total number of appointments, and treatment dates at similar seasons wherever possible. No one was selected for a control who had ever had an appointment canceled by his therapist up to, and including, the sessions investigated here.

All *S*s had at least four appointments before, and four after, the matched appointment. The attendance or absence of each *S* was recorded for each of these eight appointments.

Results

Table 1 presents the number of cancellations by both groups on the eight appointments encompassing the matched appointment. The term cancellation is used to include any absence from an appointment, whether *S* actually canceled it, or merely failed to keep it.

Table 1

Number of Patient Cancellations before and after the Matched Appointment

Appointment	Group *E*	Group *C*
Before		
4th	6	11
3rd	14	16
2nd	7	11
1st	9	17
Matched appointment	—	0
After		
1st	6	17
2nd	11	12
3rd	13	10
4th	10	10

The data in Table 1 do not support the hypothesis that a patient will cancel his next appointment after his therapist cancels. There is, in fact, a decrease in the number of patient cancellations on that appointment. Only six *S*s in group *E* canceled the appointment which followed the appointment canceled by their therapists. This figure is only 14.9 per cent of the 47 cases of therapist cancellations. In other words, about 85 per cent of the patients attended the first appointment following the one that the therapist suddenly canceled. Patients seem to be more anxious to see their therapists after missing an appointment than to try to "get even" by canceling. The degree of association between therapist and patient cancellations and attendances as determined by the phi coefficient is —.88. This value indicates that there is a marked negative association between these variables.

Confirmation of one aspect of the negative relationship discovered here may be found in a comparison of the number of cancellations made by each of the two groups on the first appointment after the matched appointment. Cancellation by the therapists is followed by high patient attendance, and attendance by all the therapists is followed by high patient cancellation. The chi-square test was applied to find the significance of these data. A fourfold contingency table was constructed showing the association of patient attendance and cancellation in the two groups. The obtained chi square of 6.96 exceeds the .01 level of confidence for 1 *df*. The two groups can therefore be said to differ significantly with respect to these cancellations. There were fewer cancellations on the first appointment after the matched appointment in group *E* than in group *C*, and this difference is not the result of sampling variations.

Discussion

The cancellation frequencies in this study suggest trends which could be validated only by the selection of new groups. Perhaps the greater frequency of "before" cancellations in group *C* may be explained by the fact that the *S*s for group *C* were selected because they had kept a certain appointment. This selection may have increased the chances of getting *S*s with a high number of cancellations just prior to the matched appointment. Cancellations by *S*s in group *C* are consistently higher than those by *S*s in group *E* on the four "before" appointments, and are strikingly so on the first appointment before the matched one. Where 100 per cent attendance occurs, a fairly high amount of cancellation has gone before. Whether canceling is done by the therapist or by the patient may not be important—at least

for the attendance record.

Matched pairs of *S*s were used to determine whether or not the two groups were selected from the same population, and did not differ with respect to their frequency of cancellation on the four appointments preceding the matched one. Subtracting the number of "before" cancellations by each *S* in group *C* from those of the paired *S* in group *E* yielded an M_{diff} of —.40 with an SE_{diff} of .19. A *t* value of 2.16 was found, and with 47 *df*, *p* is between .02 and .05, indicating that this difference must be regarded as one of doubtful significance.

Testing the groups for homogeneity in terms of their frequency of cancellations over the whole period studied, however, indicates that the groups were drawn from a single population. With matched pairs of *S*s, difference scores were obtained by subtracting the number of cancellations in group *C* from the number in group *E* for the four appointments before the matched appointment and also for the four appointments after it. The difference between the number of "before" and "after" cancellations was then found for each pair. The M_{diff} was .13 with SE_{diff} .24. The *t* value of .54 indicates that the M_{diff} is not significant ($p = .60$, $df = 47$).

Summary

The purpose of this study was to investigate the hypothesis that patients will cancel the appointment which follows an appointment suddenly canceled by their therapists. Appointment records were used to obtain a count of the cancellations of patients for the four appointments before and four after the matched appointment selected for observation. The *S*s were 47 clinic patients whose therapists had suddenly canceled an appointment with them, and 47 patients whose therapists had kept an appointment which they (the patients) also kept. The records revealed that only six of the 47 *S*s whose therapists were absent for an appointment canceled their next appointment. Thus the main hypothesis was not verified. A statistical analysis of the data revealed a significant difference between the groups with respect to the number of cancellations on the appointment immediately following the matched appointment. Incidental findings support the possibility that high cancellation frequency is followed by high attendance. In the control group, the highest number of cancellations occurred not only immediately after, but also immediately before, the matched, i.e., kept, appointment. Further studies on new samples are needed to confirm discoveries made in this tally.

Received August 10, 1953.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

The Comparative Effectiveness of Individual Psychotherapy, Group Psychotherapy, and a Combination of These Methods[1]

George O. Baehr
VA Hospital, Danville, Illinois

A large body of psychotherapists [8] regards with skepticism the notion that group methods can be effective in the treatment of emotional disorders. Yet, psychotherapists who have applied group methods systematically, notably Slavson [9], Klapman [5], and Dreikurs [2], have contended that group psychotherapy is not only effective but that it affords certain benefits not obtainable through individual therapy. According to this contention, individual and group psychotherapy can be expected to supplement one another to the greater benefit of the patient. The present investigation was designed to test the hypothesis that, within the limits to be set forth later, treating patients by a combination of group and individual psychotherapy is more effective than treating them by one method alone.

Procedure

Effectiveness, as a measure of therapy, does not have a universally accepted criterion. Effectiveness of therapy might be evaluated in terms of the patient's tendency to act out antisocial impulses, it might be evaluated with respect to his ability to plan realistically for the future, or again, it might be evaluated in terms of the patient's ability to hold a job and earn a living, etc. The criterion of therapeutic effectiveness employed in the present study is *discontentment*. Discontentment was chosen as the criterion because it is the central problem of the patient who voluntarily applies for treatment, because it is regarded by the investigator as the focal expression in consciousness of all intra- and interpersonal conflict and malfunction, and because it can be represented in quantitative terms. Conceptually defined, discontentment is the sense of ill-being coupled with dissatisfaction. Operationally, it is defined by a scale which was especially designed to measure the subject's discontentment level. This scale, as viewed by the subject, allows him "to show how much he is bothered" by each of his symptoms.

The Discontentment Scale [1, pp. 44-57] consists of 230 statements describing potentially discontentment-arousing experiences involving persons, the supernatural, sex, body functions, thought processes, the senses, the ego ideal, the superego, etc. The subject is required to place each statement which is true about him on an eight-interval scale to indicate the intensity with which each experience troubled him during the past week.

The average value that a subject assigns to the statements is taken as the best single index of his level of discontentment during the week immediately preceding the examination. It needs to be emphasized that the Discontentment Scale yields an index of the intensity of experience. The index is not based on the number of experiences, as in the MMPI [3], nor is it based on the frequency with which experiences occur, as in the second section of the Inventory of Psychological Problems developed by the Air Force during World War II [4, p. 71].

By means of a suitable formula, the scale values are converted to discontentment indices ranging from 0 to 100. For a sample of 86 newly hospitalized patients, the 24-hour test-retest reliability of the discontentment index was found to be +.84; whereas for a sample of 28 patients about to be discharged after the completion of their treatment, the 6-hour test-

[1]From the Veterans Administration Hospital, Danville, Illinois. The data for this study were collected at the Veterans Administration Hospital, Hines, Illinois.

retest reliability was found to be +.96. On two samples ($N = 10$; $N = 11$), the discontentment index was found to have an average product-moment correlation of +.47 with the ward physicians' estimates of the patients' need for treatment.

Each subject was given the Discontentment Scale at the beginning and end of his period of hospitalization. The effectiveness of psychotherapy for each individual was determined by calculating his relative movement index using the formula:

$$m_r = \frac{100\,(d_i - d_t)}{d_i}$$

where
m_r is the relative movement in the discontentment level,
d_i is the discontentment level before therapy (the initial discontentment level), and
d_t is the discontentment level after therapy (the terminal discontentment level).

The relative movement index has been found to have a product-moment correlation of +.66 ($N = 66$) with the patients' own evaluation of the amount of benefit derived from their period of hospitalization.

In the present investigation, the experimental population consisted of 66 World War II hospitalized, male veterans ranging in age from 20 to 58 years, with the median age at 30. They met the following criteria: voluntary admission for treatment, voluntary discontinuance of treatment and discharge from the hospital, a diagnosis of psychopathology, and legal competence. These criteria, rather than narrow diagnostic categories, were made the basis of selection because of the well-known unreliability of the latter [4, p. 85] and because the criterion of psychotherapeutic effectiveness employed was felt to be appropriate only for voluntary patients. Diagnostically, the experimental population consisted of 2 early schizophrenics, 1 manic-depressive of the depressive type, 40 psychoneurotics, and 23 patients suffering from character and behavior disorders. The subjects were treated with nondirectively oriented individual and group psychotherapy in various proportions and divided into the following categories:

A. Those treated predominantly by group therapy

B. Those treated predominantly by individual therapy

C. Those treated by an approximately equal amount of group and individual psychotherapy

D. Those treated predominantly by one method of psychotherapy, i.e., categories *A* and *B* combined.

Division of the total experimental population into these categories was accomplished with the aid of the formula:

$$R = \frac{G - I}{G + I}$$

where
G represents the hours of group therapy,
I represents the hours of individual therapy, and
R represents the ratio of the difference between group and individual therapy hours to the total number of hours in psychotherapy.

R has a negative value if individual therapy exceeds group therapy, has a positive value if group therapy exceeds individual therapy, and is equal to zero if the amount of time spent in each is the same.

Two different procedures were used to test the hypothesis.

1. In procedure 1 the 66 experimental patients were divided into three categories. The 17 patients (approximately 25%) who had a predominance of group therapy ($R \approx +1$) formed category *A*; the 16 patients (approximately 25%) who had a predominance of individual therapy ($R \approx -1$) formed category *B*; and the 33 patients (50%) who had an approximately equal amount of group and individual therapy ($R \approx 0$) formed category *C*. Each subject's relative movement index was calculated separately and the mean of these measures was computed for each category. These means indicate the effectiveness of the particular methods of therapy applied to the categories. The *t* test was employed to determine the significance of the difference between the means. Since the purpose of the investigation was to determine the functional relationship between the methods of therapy and movement in the discontentment level, precautions had to be taken to ensure that the categories were not significantly different with respect to other variables which might influence movement in the discontentment level. These variables are the number of therapy hours, the duration of hospitalization, and the initial discontentment level. In no instance did the categories differ on these variables at even the .10 level of confidence.

2. In procedure 2 the 66 experimental pa-

tients were divided into two categories. Category D consisted of those 33 patients whose R value indicated a predominance of a single method of therapy ($|R| \approx 1$); and category C, as in procedure 1, consisted of those 33 patients who had an approximately equal amount of group and individual therapy ($|R| \approx 0$). In a preliminary check it was found that these two categories differed, to a statistically significant degree, with respect to the mean number of therapy hours. In order to equate the two categories, the three patients in category D who had the smallest number of therapy hours were eliminated and the three patients in category C who had the largest number of therapy hours were eliminated [6, p. 448]. The two categories were then comparable on all control variables (cf. Table 2).

Results

Procedure 1, which divided the experimental population into three categories, led to the finding that the category treated with both group and individual therapy had the highest relative movement index whereas the category treated predominantly with group therapy had the lowest movement index. These findings are in keeping with the hypothesis of the investigation and suggest that the methods rank with respect to therapeutic effectiveness as follows: group-plus-individual therapy, individual therapy alone, and group therapy alone (the least effective). However, the findings are at best only suggestive since the only t that was significant even at the .10 level of confidence was the one between the combination method and the group method. The results for procedure 1 are shown in Table 1.

The results from procedure 2, which divided the experimental population into two categories, are given in Table 2. It will be seen from the table that there is a significant difference (at the .02 level of confidence) between the two categories with respect to the relative movement index. This indicates that the group-plus-individual method was found to be more effective than one method of psychotherapy and thus confirms the hypothesis of the investigation.

Table 2 also gives the statistical data on several other variables. Within the limits of the experimental design, three of these variables, viz., the total number of therapy hours ($G + I$), the duration of hospitalization (H), and the initial discontentment level (d_i) could inadvertently have functioned as independent variables, i.e., could have functioned to contribute to the difference between the relative movement indices. The statistical data in Table 2 indicate that this possibility is unlikely, since there are no significant differences between the two categories on these variables.

Table 1

Comparison of Group Psychotherapy, Individual Psychotherapy, and the Combination of These Methods

(The relative movement index [m_r] is the measure of therapeutic effectiveness.)

Item	Category (by R value) A: Predominance of group psychotherapy	B: Predominance of individual psychotherapy	C: Individual and group therapy approximately equal
N	17	16	33
Mean m_r	+5.35	+10.06	+23.54
t for difference between means	0.419	1.652	
		1.969*	

* Significant at the .10 level of confidence.

Table 2

Comparison of the Combination Method of Group and Individual Psychotherapy with the Situation in Which One Method of Psychotherapy Predominates

(The relative movement index is the measure of therapeutic effectiveness. N = 30 for each category.)

	Variable	Category (by |R| value) C: Group and individual psychotherapy Mean	C SD	D: One method of therapy predominating Mean	D SD	Diff. betw. means	t
m_r	Relative movement index	+25.93	25.22	+ 6.87	31.82	19.06	2.529*
G+I	Hours of psychotherapy	28.63	13.82	20.27	11.81	8.36	.997
H	Duration of hospitalization in days	61.87	28.86	57.73	17.29	4.14	.662
d_i	Initial discontentment index	53.17	17.38	52.03	17.20	1.14	.250
d_t	Terminal discontentment index	37.83	15.35	45.10	14.14	7.27	1.875†

* Significant at the .02 level of confidence.

† Significant at the .10 level of confidence.

In addition to the significant difference between the two categories with respect to the relative movement index, the table also shows a difference (significant at the .10 level) in the terminal discontentment index, the category treated by the combination method having the lower terminal index. The finding of this relationship is consistent with the fact that, while the initial discontentment levels of the two categories were equal, the relative movement in the discontentment level of the category treated by the combination method of therapy was significantly greater.

Discussion

It is important that the limitations of this investigation be made explicit.

1. A completely adequate evaluation of therapeutic effectiveness would take into account not only the degree and rapidity with which therapeutic changes are brought about but also the persistence of these changes after termination of therapy. The present study deals only with the degree of change over the course of therapy.

2. The problem of the therapeutic goal, i.e., the criterion for therapeutic effectiveness, has many solutions. This research employed discontentment and the relative movement in the patient's discontentment level as the criterion and the measure of the effectiveness. [Cf. 1, pp. 9-43, for a detailed discussion on the use of this criterion.]

3. The findings are limited to voluntary patients with psychopathology. They should not be generalized, without further evidence, to involuntary patients.

4. Although secondary sources of training and experience, and differences in the personalities of the psychotherapists, are likely to have introduced a certain amount of eclecticism in therapeutic approach, the therapeutic atmosphere was, nevertheless, fairly uniform from patient to patient. The 12 persons who administered the psychotherapy were all clinical psychologists trained in the nondirective methods of Rogers [7]. The therapy emphasized personal problems and emotional expression. In addition, in group therapy, group interaction was fostered. Lectures on psychology and other topics, and the giving of deep interpretations and advice were judiciously avoided. The findings of this investigation must be understood in the light of this setting.

5. In the present research, the expression "combined method of therapy" has meant group therapy followed by individual therapy. Without further evidence, we cannot generalize to courses of psychotherapy in which this sequence is reversed or where group and individual psychotherapy are carried on simultaneously.

Within these limitations, the findings of the present investigation permit the inference that the combination of group and individual psychotherapy is more effective than one method alone. They also suggest that, with respect to their effectiveness, the three methods rank, from most to least effective, as follows: the combination of methods, individual psychotherapy alone, group psychotherapy alone.

Various writers who have had experience with both group and individual psychotherapy have recorded clinical observations that have implications for the relative effectiveness of these methods used singly and in combination. At the International Congress on Psychiatry held in Paris in 1950, Dreikurs [2, p. 5] stated that, in 1929, he quite accidentally stumbled upon the discovery that treating patients in groups affords therapeutic benefit not yielded by individual therapy alone. Since that time, he has used group therapy as an adjunct of individual therapy, clearly implying that he considers the combination to be more beneficial, criteria not stated, than either method alone.

Klapman implies a similar ranking for effectiveness when he writes, "as catalyst, furthermore, group therapy will be found efficacious in speeding the dissolution of resistances. When in individual interview the patient ceases to produce any material, 'shuttling' him into a group session may not only help in overcoming resistance but even be of diagnostic importance" [5, p. 89]. Nolan D. C. Lewis, in his foreword to a volume edited by Slavson, expressed himself more cautiously, writing, "Experts in Group Therapy do not claim that it is a substitute for individual treatment, but feel that it is indicated in special, carefully chosen cases, with a proper assignment as far as the composition of the group is concerned" [9, p. 10].

Although the experimental results from procedure 1 are not conclusive in this respect, they are consistent with the clinical observations of these writers, insofar as the methods were found to rank in the order: combined therapy, individual therapy alone, group therapy alone (least effective).

The experimental results from procedure 2 strongly suggest that the combination method of therapy is more effective than either the group or the individual method alone. A superficial explanation may be inferred from the quotations. If there are benefits to be derived from group therapy that cannot be provided by individual therapy under certain circumstances and if the converse is also true, then, barring the possibility of interferences, we would expect their combination to produce greater benefits than either alone.

Although the present study and Dreikurs' observation come to the same conclusions regarding the relative effectiveness of the three methods of therapy considered, our experimental conditions provided an essentially nondirective type of therapy (client-centered psychotherapy), while Dreikurs used a directive type (individual psychology). Also, Dreikurs reversed the sequence of the therapies, applying first individual, then group psychotherapy. The evidence suggests that there is something about combining group and individual therapy which makes for greater effectiveness, independent of the order of application and the type of therapy used. How far the type of therapy can differ from those mentioned here and still conform with this generalization is a problem for further research. Accepting the high probability that group therapy has the capacity to augment individual therapy, the further question is raised for investigation: What are the particular therapeutic ingredients that each therapy supplies?

Summary

The investigation was designed to test the hypothesis that, within the limits described, treating voluntary patients by a combination of group and individual psychotherapy is more effective than treating them by one method alone. The findings strongly suggest that while all three methods produce beneficial results, the combined method is superior to one method alone. They also suggest that the methods rank themselves, from most to least effective, in the following order: group-plus-individual psychotherapy, individual therapy, group therapy.

Received October 15, 1953.

References

1. Baehr, G. O. The evaluation of psychotherapeutic results with special reference to self-referred patients. Unpublished doctor's dissertation, Univer. of Chicago, 1951.
2. Dreikurs, R. R. Psychotherapie de groupe. Paper read at the International Congress on Psychiatry, Paris, 1950. Distributed in mimeographed form.
3. Hathaway, S. R., & McKinley, J. C. *Manual for the Minnesota Multiphasic Personality Inventory.* New York: Psychological Corp., 1943.
4. Heathers, G. L. Personality and adjustment studies. In S. W. Bijou (Ed.), *The psychological program in AAF convalescent hospitals.* Washington: U. S. Government Printing Office, 1947. Pp. 59-85.
5. Klapman, J. W. *Group psychotherapy, theory and practice.* New York: Grune & Stratton, 1946.
6. Peters, C. C., & Van Voorhis, W. R. *Statistical procedures and their mathematical bases.* New York: McGraw-Hill, 1940.
7. Rogers, C. R. *Counseling and psychotherapy.* Boston: Houghton Mifflin, 1942.
8. Shaskan, D. A. Must individual and group psychotherapy be opposed? *Amer. J. Orthopsychiat.*, 1947, 17, 290-292.
9. Slavson, S. R. (Ed.) *The practice of group therapy.* New York: International Universities Press, 1947.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Body Size and Body-Cathexis

Sidney M. Jourard and Paul F. Secord
Emory University

In an earlier paper [1], the writers described a method for the measurement of body-cathexis, and reported that body-cathexis — the degree and direction of feelings toward one's own body — is related to anxiety, insecurity, and feelings toward the self. The present study tests whether the actual measured size of body characteristics is also a correlate of body-cathexis.

This expectation arose from consideration of the popular stereotype of the male ideal, exemplified by the tall, broad-shouldered, muscular hero in fiction and the movies.

Method

Subjects. Sixty-two male undergraduates from the introductory psychology course at Emory University served as *S*s.

Procedure. Each *S* filled in a body-cathexis questionnaire. This form, described in detail elsewhere [1], but since modified slightly, consists of a listing of 40 aspects of the body. The respondent indicated the nature of his feelings toward each item in accordance with the following five-point scale:

1. Have strong positive feelings.
2. Have moderate positive feelings.
3. Have no feeling one way or the other.
4. Have moderate negative feelings.
5. Have strong negative feelings.

Only a few *actual size* measurements related to the "masculine-ideal type" were feasible. After completion of the cathexis scale, and without prior knowledge, measurements were taken from each *S* (using a tape measure and scales) of his height, weight, shoulder width, chest circumference in relaxed position, expanded chest circumference, biceps circumference with the arm hanging loosely, and biceps when contracted.

Correlations were computed between these measurements and the corresponding cathexis ratings on height, weight, shoulders, chest, and muscular strength.

Table 1
Correlations between Body Measurements and Body-Cathexis

Body-cathexis	Body measurements	r*
Height	Height	−.55
Weight	Weight	−.04†
Shoulders	Width of shoulders	−.36
Chest	Chest (relaxed)	−.33
Chest	Chest (expanded)	−.37
Muscular strength	Biceps (relaxed)	−.52
Muscular strength	Biceps (contracted)	−.54

* $r = .33$ at .01 level for $df = 60$.

† eta for weight and cathexis = .19; eta for $\frac{\text{weight}}{\text{height}}$ and weight-cathexis = .34 ($F = 1.09$, $p > .05$).

Results

As may be seen in Table 1, correlations significant beyond the .01 level were obtained for all variables except weight. The r's signify that a significant portion of the variance in cathexis scores is accounted for by variation in actual size. The direction of the r's indicates that large size is associated with strong, positive feelings toward the respective body part, and small size with weak or negative feelings.

Discussion

With the exception of body weight, the findings reveal that cathexis for body characteristics pertaining to "masculinity" is related to the size of relevant body parts. Large size is apparently a desired quality among males, and its presence or absence leads to contrasting feelings toward related aspects of the male body.

The absence of a significant relationship between weight and cathexis (neither r nor η reached significance) indicates that there is no simple relationship between these two variables. The η between weight/height ratio and weight-cathexis was higher but still insignificant.

Received September 24, 1953.

Reference

1. Secord, P. F., & Jourard, S. M. The appraisal of body-cathexis: body-cathexis and the self. *J. consult. Psychol.*, 1953, **17**, 343-347.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Validation of Rorschach Criteria for Predicting Duration of Therapy [1]

Robert G. Gibby, Bernard A. Stotsky
VA Mental Hygiene Clinic, Detroit, Michigan
E. Wesley Hiler, and Daniel R. Miller
University of Michigan

Within the past two years, there have appeared in this journal four articles that dealt with the use of the Rorschach test to predict duration of therapy. Although they used similar techniques and populations, these studies contradicted each other in a number of major aspects. The present investigation was designed to treat some of these areas of contradiction with a view to discovering the principle which accounted for the discrepancy. In order to elucidate the present design, we will describe the primary issues in terms of the findings of the previous investigators.

Possibility of prediction. The first article, by Rogers, Knauss, and Hammond [5], reported no significant difference on 99 Rorschach signs between patients at a Veterans Administration Mental Hygiene Clinic who stayed in therapy for four or less sessions and those who remained five times or more. However, Kotkov and Meadow [4], studying veterans engaged in group psychotherapy, obtained by the use of discriminant function a weighted combination of three scores, *R*, *D%*, and *FC — CF*, which discriminated with 69 per cent accuracy between patients with fewer than nine sessions and those who stayed in nine sessions or more. In validating their formula on a second sample of patients in individual therapy, Kotkov and Meadow again found that they could discriminate with 69 per cent accuracy.

Stimulated by the success of Kotkov and Meadow, Auld and Eron [1] tried the same formula on a group of 33 patients in a general hospital with no success. The accuracy of discrimination between patients who stayed in and those who left before nine sessions was 52 per cent, which is only 2 per cent from chance. The number of responses alone discriminated significantly. With their cutting point of 19, the accuracy was 67 per cent.

The present writers did an initial study shortly after the negative results of Rogers, Knauss, and Hammond appeared. In an attempt to determine whether the Rorschach test could predict duration of therapy for extreme groups, we attempted to discriminate between those who stayed in therapy for five sessions or less and those who remained for twenty or more. We did not attempt to establish a formula, but were interested solely in the specific Rorschach signs which *discriminated* between these two extremes. We were encouraged by the finding that differences between those who terminated and those who remained were very significant for a large series of scores, although a fair proportion of them were correlated significantly with the number of responses. Because we were partially successful in that endeavor, we are following it up with the present study.

Sample and Method

Subjects were selected from among patients at the Veterans Administration Mental Hygiene Clinic at Detroit, Michigan. The 269 patients in the study were predominantly neurotics, with some character disorders and ambulatory psychotics included. The socioeconomic level was predominantly lower and lower middle. The average amount of education was eleven years; the average age, 30. Most of the patients were white, only 10 per cent being Negro. Excluded from both groups were disoriented psychotics, organics, females, and patients who were forced to terminate for reasons beyond their control, such as shift in

[1] From the VA Regional Office, Detroit, Michigan.

residence, etc. The precedent established by Kotkov and Meadow was followed. By the method of discrimination function, a formula that maximally differentiated between the 33 continuing patients and the 51 terminators was derived. Then the obtained formula was validated with a sample of 75 continuing subjects and 110 terminators.

Cutting point. In our previous article we raised the possibility that the negative results of Rogers, Knauss, and Hammond might be attributable to their cutting point of five therapeutic sessions. It did not seem reasonable to us to assume that a person who stays in therapy for six or seven interviews can be regarded as having gained significantly more from therapy than a person who stays for only four sessions. Kotkov and Meadow, who were partially successful in making predictions, used a cutting point of nine sessions; so did Auld and Eron who, although they did not discriminate significantly in terms of Kotkov and Meadow's formula, did succeed with *R* alone. Some of our most significant results were based on the cutting points of five and twenty sessions.

We decided to include in the continuing group only those patients who had been in therapy at least six months and, preferably, were still in treatment at the time that the data were collected. Inclusion in the terminating group, then, required the patient to have been in therapy less than six months. The cutting point derived from this criterion turned out to be 19 sessions or less.

Although 19 sessions may seem high as a point of division for the two groups, it was felt that, in attempting to predict continuation in therapy, it would be necessary to account for every individual who left treatment against the therapist's advice. This made it necessary that the sample include an intermediate group of patients who terminated therapy between their sixth and nineteenth sessions. Presumably, this group was more motivated than were the patients who dropped out between the first and fifth sessions. In theory, one might expect this group of patients to resemble the continuing group more than did the group which dropped out of treatment early. Analysis of the Rorschach data showed this not to be true. The patients in this group were more similar to those in the one-to-five session group than to patients continuing in therapy. For instance, the average number of Rorschach responses for this group was approximately 20; for the early "drop" group, also 20; for the "stay" group, 31. In view of the test performance of the intermediate group, our decision to place them with the terminators appears to have been a wise one. In our previous article we suggested that choice of a cutting point could influence the results considerably. For instance, by choosing five as a cutting point, Rogers, Knauss, and Hammond may well have misclassified a considerable number of "drop" patients. The same issue arises with the samples of Kotkov and Meadow and those of Auld and Eron. If discrepancies should arise between the findings of the present study and those of the previous ones, one possible explanation could be the choice of different cutting points.

Results

1. *The prediction formula.* On the first sample, the formula obtained by the method of discriminant analysis that best differentiated between terminating and remaining patients was

$$Y = .00049\,R + .00212\ \text{Total}\,K + .00010m.$$

The term Total *K* refers to the sum of all the *K* responses, none of which was weighted in any way. We applied the formula to our first sample. By using a cutting point of .02, we were able to differentiate between those who terminated and those who remained with an accuracy of 68 per cent. On applying the formula to the second sample, we were able to

Table 1

Frequency Distribution of *Y* Scores for Remaining and Terminating Patients

Y-score class intervals	Sample I		Sample II	
	R	*T*	*R*	*T*
.06000 – up	1	0	1	0
.05000 – .05999	2	0	2	1
.04000 – .04999	0	0	0	0
.03000 – .03999	5	1	4	2
.02000 – .02999	6	7	21	11
.01000 – .01999	14	19	35	49
.00000 – .00999	5	24	12	47
Total *N*	33	51	75	110

predict accurately 67 per cent of the time those who remained and those who terminated.

While 67 per cent accuracy is not sufficiently high to make this formula useful to the practicing clinician, our results nevertheless corroborate the findings of Kotkov and Meadow and of Auld and Eron that some sign could be found which would discriminate significantly between continuing therapy and terminating.

It should be noted from Table 1 that the formula is much more successful in predicting termination than continuance of therapy. Approximately 84 per cent of the terminators are correctly classified in the first group. The prediction of terminators in the second sample is approximately 87 per cent accurate. We have received a communication from Meadow in which he points out that the Kotkov and Meadow formula is only successful in making predictions for the group of patients who do not continue in psychotherapy. While it yielded more incorrect predictions for the continuing group in his study, the accuracy of locating terminators was 88 per cent. He also makes the cogent point that the negative results of Auld and Eron's experiment may have resulted from the fact that they used 21 patients for their continuing group and only 12 for their noncontinuing group. We then wrote to Dr. Auld, and he supplied us with the scores of his terminators. These discriminated correctly in 8 of the 12 cases. The 67 per cent accuracy, while not as high as that of Kotkov and Meadow or as high as ours, was still much higher than the over-all accuracy for both continuers and terminators.

It may be of some interest to remark on the performance of the intermediate group of 53 patients. Using .02 as a point of division, 83 per cent were classified correctly as terminators. This group was almost indistinguishable from the early terminators, of whom 88 per cent were classified correctly. These findings support the categorization of the intermediate group as terminators.

The Kotkov-Meadow formula was applied to the 185 cases in our validation sample. Patients with a *Y* score of .01132 were assigned to the "continuing" group and those with lower scores to the "discontinuing" group. Sixty-one per cent of the patients were classified correctly. This is slightly better than chance but considerably below the 69 per cent accuracy obtained in the studies of Kotkov and Meadow.

2. *Productivity* (R) *as a predictor.* A common finding in all these studies was that the total number of responses (*R*) was consistently associated with continuance in therapy. In our sample both Total *K* and *m* correlated over .60 with *R*. Therefore, we decided to test the ability of *R* by itself to predict continuation. By inspection we found 35 responses to be the best point of division between remainers and terminators in Sample I.

Forty-eight terminators and 21 remainers gave fewer than 35 responses; only three terminators and 12 remainers gave 35 or more. Accuracy of prediction was 71 per cent in Sample I. In the validation sample 102 terminators and 51 remainers gave fewer than 35 responses, while 8 terminators and 24 remainers gave 35 or more, for a total accuracy of 68 per cent. Considering the two samples together, 69 per cent of the patients were classified correctly by use of *R* alone. These findings would tend to suggest that *R* by itself can predict as well as any group of Rorschach variables, provided that a suitable cutting point is chosen. This cutting point, however, may vary from clinic to clinic. Auld and Eron found that an *R* greater than 18 predicted quite well for their sample.

The question inevitably arises as to how much better than chance were these predictions. For the entire sample, chi square based on a fourfold table was greater than 30, with 35 responses as a point of division. This value was significant beyond the .001 level. We may consequently conclude that the number of cases correctly classified was better than chance.

What would happen if one predicted termination for every case? The accuracy of prediction would be roughly 60 per cent, since that is about the proportion of terminators in the sample. This prediction is only 9 per cent poorer than the best Rorschach variable. Although the difference is significant, the increment in prediction resulting from *R* is not very great.

3. *Productivity and intelligence.* In order to explore the reason why number of responses discriminated between continuers and termin-

ators, Auld and Eron first correlated *R* with continuance. The obtained coefficient was +.55. However, when the verbal scale of the Wechsler-Bellevue was held constant, the correlation between Rorschach productivity and continuation dropped to +.07. In fact, the IQ scores gave better predictions of continuance than the number of Rorschach responses, with a biserial correlation of +.71, which was significant at the .05 level.

The authors concluded that the Wechsler-Bellevue verbal scale was more useful than the Rorschach for predicting continuance in therapy. Before accepting as valid such use of the verbal scale, they felt that their finding based on 23 patients should be supported by evidence from cross-validation studies.

This the present authors attempted to do. In the first sample there were 79 Wechsler verbal scales available, of which 37 were of remainers and 42 of terminators. The best cutting point was empirically determined at 100. A Verbal IQ score of 100 and up was considered a favorable sign; below 100, unfavorable. Fifty out of 79 patients were correctly classified, a 63 per cent accuracy. On the validation sample, 54 out of 88 were categorized correctly, a 61 per cent accuracy.

Inasmuch as our findings gave only qualified support to those of Auld and Eron, we decided to study the relationship between *R* and Verbal IQ in our sample. For 165 cases where there were both Wechsler verbal scales and Rorschachs available, the correlation between *R* and Verbal IQ was +.24. The biserial correlation between number of responses and continuance was + .38; between Verbal IQ and continuance, +.40. With *R* partialled out, Verbal IQ correlated +.35 with the criterion. With Verbal IQ partialled out, *R* correlated +.32 with continuance. In our sample the relationship between *R* and continuance was not entirely accounted for by the correlation with Verbal IQ since *R* remained significant with the verbal score partialled out. The multiple correlation of *R* and Verbal IQ with continuance was +.50. Our findings do not support the contention of Auld and Eron that the relationship between number of responses and continuance in therapy only demonstrates that more intelligent patients are more likely to continue.

4. *Social class as a predictor.* In a personal communication to one of the authors, Auld suggested that social class may be related to persistence in therapy. The present authors undertook to test the hypothesis that socioeconomic class was related to staying in therapy, the prediction that middle class patients would stay and lower class patients and unemployed patients would leave treatment. The sample was broken down into five classes:

1. Upper middle—low managerial, highly skilled clerical, and professional workers.
2. Lower middle — semiskilled and unskilled clerical and semiprofessional workers.
3. Upper lower — skilled manual workers.
4. Lower lower — unskilled manual workers.
5. Unemployed.

For the first sample there were 76 classifiable subjects of whom 47 were classified correctly, most of them in the terminating group. Sixty-two per cent accuracy of prediction was achieved; the chi square was nonsignificant, though positive. For the validation sample, 111 out of 177, 63 per cent, were categorized correctly. The chi square of 8.63 was significant at the .01 level. Since in both samples the trend was in a positive direction — with middle class people tending to remain in and lower class people to leave therapy — the hypothesis that class is related to continuance in therapy is supported in this study.

5. *Interaction between patients and therapists.* A factor not considered in previous investigations which might exert considerable limitation upon the predictive efficiency of the Rorschach test is the skill of the therapists and the nature of their resultant interactions with the patients. We were fortunate to be able to obtain data that not only bear upon this interaction but may account in part for some of the discrepancy in the results of other investigators.

Hiler [3], and Hiler and E. L. Kelly in a study as yet unpublished, did an analysis of patient-therapist interaction in order to determine whether therapists differed in respect to the types of patients they were able to keep in treatment. The number of Rorschach responses, which was assumed to have as components productivity and motivation for treatment, was used to characterize the patients.

Their findings are reported briefly below:

1. It was found that whether or not unproductive and poorly motivated patients remained in treatment was in part a function of the therapist. Although, for some therapists, terminators did not differ significantly from remainers on *R*, for most therapists they did.

2. Patients remaining in treatment were similar in respect to productivity on the Rorschach whether the therapist was a psychiatrist, psychologist, or social worker. The same was true for patients terminating.

3. Productivity on the Rorschach was not related to continuance in therapy with the female therapists employed in this study, but was associated with continuance for patients seen by male therapists. The female therapists were able to keep in treatment patients giving few Rorschach responses, but tended to lose proportionately more productive (high *R*) patients.

4. Therapists rated as warm and friendly were able to keep in treatment more of the poorly motivated, unproductive patients (low *R*) than were therapists rated as cold and unfriendly.

5. Therapists judged to be most competent at analytically oriented therapy lost proportionately fewer productive patients than did the less competent ones.

6. There was no significant difference between therapists rated high on passivity and those rated low on passivity with regard to the proportion of nonproductive patients remaining in treatment.

Discussion

Our study confirms the findings of previous studies that the total number of Rorschach responses (*R*) is significantly related to continuance in therapy. It was also found that *R* by itself predicted as well for both samples as did the discriminant function using *R*, Total *K*, and *m*. Both *R* and the discriminant function derived from our first sample predicted better than the Kotkov and Meadow formula which had been derived from a group of similar patients.

The findings from these studies, together with significant differences between groups on Verbal IQ and social class, suggest a number of tentative conclusions: (*a*) Patients who remained were, in general, more intelligent and more productive. (*b*) They appeared to possess the ambition and drive associated with the middle class and upwardly mobile lower class persons in American society. (*c*) They showed a greater awareness of the psychological factors responsible for their problems and, although they experienced greater anxiety, they were not incapacitated by their anxiety. (*d*) Patients who remained in treatment tended, to a greater extent than the terminators, to have both the capacities and the motivational state required to make effective use of the psychoanalytically oriented therapy offered at the clinic [2, 3].

Auld and Eron correctly stressed the importance of cross validation of any preliminary findings. This we have attempted to do by validating the findings from our first sample on a second sample more than twice as large as the original. In our study the requirement of cross validation was met since the formula worked as well in the second sample as it had in the first.

Auld and Eron also pointed out the need for validity generalization; that is, applying the data to a new and different sample with different characteristics. This step would require successive validation of the findings on a wide variety of patients in different clinics. There are a number of conditions limiting application of any such formula to different populations:

1. Since the formula applies best to terminators, it may not apply as adequately in a setting where, for some reason, a larger proportion of patients stay in therapy. Our formula, derived from a group of veterans receiving treatment from a VA facility, must be assumed, until it is proved otherwise, to generalize only to similar populations. Patients attending an outpatient clinic at a private hospital where they pay fees and presumably come entirely of their own accord cannot be equated with a group of VA patients who draw pensions for neuropsychiatric disorders unless there is definite evidence to show that such groups are equivalent. While this consideration should not necessarily prevent a formula derived from one type of sample from being applied to another, the limitations of such an application must be kept in mind.

2. It has been shown that both the warmth and competence of the therapist may have an effect on the applicability of a prediction formula. It is hard to see how the operation of such factors can do anything except decrease the predictive adequacy of any formula based on test material that does not take them into account.

3. The type of therapy employed can have a direct bearing on the ability of a formula to predict. Our study was performed at a psychoanalytically oriented clinic where therapy is seen as primarily a long-term affair. At a clinic oriented to more immediate short-term goals, the style of therapy, the

choice of patient, and the criteria for determining continuation and termination would probably be considerably different from the orientation found at the Detroit clinic. In view of these different conditions, it is difficult to see how a formula could be set up to predict equally well in all situations unless these factors were all relatively unimportant in comparison with the patient's personality.

Premature termination is, then, not entirely a function of the patient's personality and motivational state, but appears to be influenced by the attitude and competence of the therapist and the type of therapy offered at a clinic. Does so-called "depth" therapy, which involves the manipulation of complex verbal concepts, make sense to individuals who are unable to handle such concepts? Would other therapeutic approaches succeed in keeping such individuals in treatment, and succeed in promoting their social adjustment? Studies such as the one reported here are needed to investigate personality differences of remainers and terminators in other settings and among therapists using different approaches to treatment.

Suggestions for Further Work in Prediction

In an unpublished study, Hiler [3] found such personal data items as social class, education, occupation, diagnosis, and initial complaints to predict continuance in therapy somewhat better than chance. The figures for these items were only slightly below those obtained by psychological test data. Not only are such data easier to collect, but combinations of such items might yield predictions exceeding the psychological tests in accuracy. Perhaps the combining of psychological test scores with personal data and data pertaining to the therapist might yield even better accuracy. For instance, a formula using scores such as Rorschach R, Wechsler Verbal IQ, social class, and initial symptoms might very well yield a high enough prediction to be of practical use to the clinician. These four types of data do not by any means exhaust the alternatives available to the researcher interested in the problem of predicting continuation in therapy. Certainly further research is needed in this area.

Summary

1. By means of discriminant analysis an attempt was made using Rorschach scores to predict whether patients will continue psychotherapy. A formula, $Y = .00049\ R + .00212$ Total $K + .00010\ m$, was derived. When applied to a sample of 84 patients from which the weights had been derived, it correctly classified 68 per cent of the subjects as continuing or terminating therapy. On a validation sample of 185 subjects, 67 per cent of the patients were correctly classified. On the same sample, the Kotkov and Meadow formula correctly predicted 61 per cent of the patients.

2. Using the total number of Rorschach responses (R) by itself, it was possible to classify correctly 71 per cent of the subjects in the first sample and 68 per cent in the second.

3. For 165 patients the relationship between R and continuance was +.38. With Verbal IQ partialled out, R continued to correlate significantly with continuance, thus contradicting the previous finding of Auld and Eron that, with Verbal IQ partialled out, R correlated insignificantly with continuance. The multiple correlation of R and Verbal IQ with continuance was +.50.

4. Social class was found to be related to continuance in therapy, with middle class subjects tending to remain in and lower class subjects to drop out of therapy.

5. The implications of these findings were discussed, especially with respect to the relative importance of such factors as type of therapy and warmth and competence of the therapist in the prediction of continuance in treatment. It was concluded that premature termination is not entirely a function of the patient's personality and motivational state. Therefore, in predicting continuance in therapy, other important factors, such as those discussed, should be considered and given their proper weight in a prediction formula.

Received August 31, 1953.

References

1. Auld, F., & Eron, L. D. The use of Rorschach scores to predict whether patients will continue psychotherapy. *J. consult. Psychol.*, 1953, **17**, 104–109.

2. Gibby, R. G., Stotsky, B. A., Miller, D. R., & Hiler, E. W. Prediction of duration of therapy from the Rorschach test. *J. consult Psychol.*, 1953, **17**, 348–354.

3. Hiler, E. W. An investigation of psychological factors associated with premature termination of psychotherapy. Unpublished doctor's dissertation, Univer. of Michigan, 1953.
4. Kotkov, B., & Meadow, A. Rorschach criteria for predicting continuation in individual psychotherapy. *J. consult. Psychol.*, 1953, **17**, 16–20.
5. Rogers, L. S., Knauss, Joanne, & Hammond, K. R. Predicting continuation in therapy by means of the Rorschach test. *J. consult. Psychol.*, 1951, **15**, 368–371.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

A Rationally Derived MMPI Scale to Measure Dependence [1]

Leslie Navran
VA Hospital, American Lake, Washington

This paper reports the development and initial use of an MMPI scale to measure dependence. An attempt to develop the scale empirically failed because experienced clinicians could not reliably rate dependence from case history summaries. A rational derivation was substituted. Sixteen judges independently specified the MMPI items that they felt to bear on dependence. They further specified the direction of response which would be made by a hypothetical dependent person who was responding frankly. Tetrachoric correlation coefficients were computed for each pair of judges, and the resulting matrix, when factor analyzed, yielded one main factor consisting of eight judges. The 157 items they collectively agreed on were screened for discriminability by internal consistency statistical procedures on a sample of 50 neuropsychiatric patients. A second very similar sample of 50 patients was used to cross validate the discriminating items so obtained. The same procedure was repeated, using Sample II first and cross validating on Sample I. Fifty-seven items survived these tests and constitute the scale at present. Reliability is .91 for the 100 patients in the derivation sample (Kuder-Richardson formula 20).

The Dependence Scale correlates .72 with *Pt*, .60 with the *D* scale, —.53 with *K*, .53 with *Pa*, and .50 with *Sc*. To date, its use with known groups has yielded these results: (*a*) 100 NP patients score significantly higher than 200 normals from the original MMPI standardization population (mean difference of 7.6 points; $p < .001$). (*b*) These normals in turn score significantly higher than 265 Stanford University graduate students (mean difference of 3.81 points; $p < .001$). (*c*) The shape of the graduate student distribution is J curved, suggesting conformity to a cultural expectation. (*d*) Within the sample of 100 NP patients, character disorders score lowest, followed by neurotics, and then schizophrenics. Differences are not significant between these groups, but nonparanoid schizophrenics score significantly higher than paranoid schizophrenics ($p < .04$; $t = 2.08$).

These results are considered promising, but not necessarily conclusive. However, the scale seems worthy of more investigation and exploratory use.

Brief Report
Received March 15, 1954.

[1] An extended report of this study may be obtained without charge from Leslie Navran, VA Hospital, American Lake, Wash., or for a fee from the American Documentation Institute. To obtain it from the latter source, order Document No. 4227 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., remitting in advance $2.00 for microfilm or $3.75 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Comparison of the Rorschach and MMPI by Means of Factor Analysis

Harold L. Williams and James F. Lawrence
Letterman Army Hospital

Although a number of research papers have been published treating separately the MMPI and the Rorschach, few comparison studies have been made. Several investigators [1, 4, 6] have explored relationships between the group forms of the two tests. These studies have not revealed striking association between them. The situation was summed up by Ellis [5] when he asserted that the MMPI does not agree well with the results of the Rorschach and other projective tests. While failure to find correlation between these widely used clinical instruments would not be convincing evidence of lack of validity for either, there seems to be an a priori basis for expecting them to correlate; that is, clinicians who use both tests may make identical predictive statements from either source (e.g., "he shows depressive tendencies").

Several separate factor analyses of Rorschach determinants and MMPI variables have now been published [10, 11, 12]. Three studies on different populations have established fair agreement as to the factorial composition of the Rorschach. Even though a general factor of productivity accounts for a high percentage of the variance in the intercorrelation matrix, the other factors found were rather similar in all three studies. The five factors found by Williams and Lawrence [10] were tentatively labeled: (*a*) productivity, (*b*) intelligence, (*c*) movement, (*d*) lack of perceptual control, and (*e*) shading.

Wheeler, Little, and Lehner [9], Tyler [7], and Welsh [8] have agreed, essentially, in describing two major and two to three minor factors among the MMPI scales. On the basis of the first two factors, Welsh[1] has developed scales which he labeled *A* and *R*. The *A* scale represents a dimension containing heavy loadings on practically all of the complaint scales, particularly *Pt* and *Sc*. Judging from item content, one finds that the scale seems related to excessive preoccupation with the psychiatrically negative aspects of the self: rumination, worry, and pessimistic attitudes. The *R* scale, based on the second major factor, is correlated positively with *D*, *K*, and *Hy*, and inversely with *Ma*. Welsh suggests that the scale may relate to expressive-repressive personality traits.

[1] Welsh, G. S. Personal communication, April, 1953.

Problem

The purpose of the present paper is to report the results of a factor analysis of the MMPI scales, certain Rorschach determinants, and the Wechsler-Bellevue Verbal IQ of 100 psychiatric patients selected at random from the files of the Clinical Psychology Section at Letterman Army Hospital. The 32 variables which were correlated can be found in Table 1.

The MMPI symbol *Es* (variable 30) refers to a scale recently developed by Barron [2] for the purpose of predicting response to psychotherapy. The scale, which is inversely related to *A*, held up rather well on cross validation. Furthermore, among normals it was found to be correlated with high ratings on a number of characteristics, such as resourcefulness, vitality, self-direction, tolerance, self-confidence, breadth of interest, emotional outgoingness, spontaneity, and intelligence. Low scorers were described as effeminate, submissive, and introspective. He suggested, therefore, that individual differences on the scale could be considered a function of ego strength.

Table 1 shows the intercorrelations computed from raw scores for the 32 variables. While the Pearson correlation coefficient was

Table 1

Intercorrelation among 31 Rorschach and MMPI Variables and Wechsler-Bellevue Verbal IQ for 100 Psychiatric Patients

Symbol Variable No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
W 1																																
D 2	19																															
d 3	17	80																														
M 4	37	75	32																													
FM 5	31	46	23	50																												
m 6	39	31	31	43	58																											
k 7	29	29	35	12	25	25																										
K 8	69	24	21	33	20	60	25																									
F 9	32	84	80	47	16	17	34	−02																								
Fc 10	28	25	10	26	22	32	32	10	00																							
c 11	55	48	37	25	26	21	30	24	46	33																						
C[1] 12	44	19	31	10	16	18	16	11	03	13	37																					
FC 13	43	40	47	47	50	55	41	30	41	40	35	30																				
CF 14	55	21	20	40	29	10	30	33	17	37	56	46	40																			
C 15	43	09	00	13	−16	−05	26	30	12	01	45	31	09	44																		
P 16	10	08	23	03	45	24	−27	−25	08	25	−01	24	05	−09	17																	
R 17	51	92	69	65	55	40	65	22	90	41	59	28	61	49	06	05																
VbIQ 18	69	25	18	30	10	00	32	29	15	12	44	12	30	22	22	−10	51															
L 19	−06	19	25	−25	00	06	18	06	19	06	20	00	−06	−06	18	18	00	−24														
F 20	−38	19	00	−25	−13	32	−37	−06	06	−12	−24	00	06	−44	18	00	00	−32	−13													
K 21	50	06	06	18	−06	00	50	32	32	−06	00	06	−24	49	18	00	18	50	21	−26												
Hs 22	−13	25	00	00	32	06	00	−06	00	−12	−25	−06	−06	13	00	−12	06	00	−06	40	−10											
D 23	−06	00	−06	06	−12	06	00	00	06	06	−25	−37	00	−32	−37	−38	06	−06	−05	41	−04	74										
Hy 24	−06	06	06	25	−06	00	06	−06	−06	−06	−06	−13	−35	00	19	−18	00	00	17	20	15	76	69									
Pd 25	−06	00	06	00	00	19	−25	18	06	00	−06	06	00	−12	00	−18	00	−24	−32	66	−16	46	51	19								
Pa 26	−06	00	00	05	00	55	−25	00	−06	−06	−10	00	−06	−25	−25	−18	−32	00	−16	76	−15	45	50	37	80							
Pt 27	−25	18	00	05	−06	00	−25	−25	06	−06	−10	00	25	−12	−25	00	00	−24	−26	67	−30	67	71	40	66	64						
Sc 28	−38	25	06	00	−25	19	−12	00	18	06	−25	00	19	−31	−23	−06	00	−40	−25	81	−25	62	62	32	77	78	92					
Ma 29	−06	00	−31	19	06	00	−06	00	−31	06	23	25	25	00	00	19	50	−16	−32	45	−10	18	00	−10	47	36	37	48				
Es 30	40	38	38	43	−06	00	40	31	38	35	−06	00	−06	50	00	35	38	39	00	−60	35	−74	−42	−48	00	−62	−87	−82	−32			
A 31	−38	00	−25	−38	−13	00	−12	−06	00	06	06	00	49	00	−25	−25	00	−31	29	67	−28	59	65	29	68	60	95	90	41	−83		
R 32	−25	−25	−06	−12	−12	−25	−12	00	38	10	−32	−06	−37	32	−06	00	−25	00	33	00	27	20	20	34	00	22	−09	−04	−58	00	−08	

Table 2

Centroid Factors

Variable No.	Symbol	I	II	III	IV	h^2
1	*W*	56	54	22	−32	75
2	*D*	76	24	−32	39	89
3	*d*	56	33	−18	18	49
4	*M*	58	30	−18	07	46
5	*FM*	43	17	−28	07	30
6	*m*	59	10	−28	−28	51
7	*k*	42	40	26	16	43
8	*K*	42	23	32	−33	44
9	*F*	61	29	−25	48	75
10	*Fc*	37	21	07	−19	22
11	*c*	51	39	04	05	42
12	*C'*	34	21	−08	−28	24
13	*FC*	63	18	−26	−14	52
14	*CF*	41	45	37	−11	52
15	*C*	22	27	24	−28	26
16	*P*	10	02	−31	−20	15
17	*R*	87	39	−20	34	106
18	*V*	30	38	26	02	30
19	*L*	00	16	20	30	16
20	*F*	27	−75	−23	18	72
21	*K*	19	30	45	10	34
22	*Hs*	35	−67	18	38	75
23	*D*	26	−69	32	41	81
24	*Hy*	24	−48	48	43	70
25	*Pd*	42	−69	17	−35	80
26	*Pa*	32	−72	−05	−18	66
27	*Pt*	37	−85	−18	18	92
28	*Sc*	39	−87	−25	−07	98
29	*Ma*	31	−32	−31	−45	50
30	*Es*	09	88	15	−21	85
31	*A*	37	−82	14	00	83
32	*R*	−16	−10	−50	10	30

used between MMPI scales, tetrachoric coefficients were used between MMPI and other variables.[2]

With few exceptions, the Rorschach determinants are not highly correlated with MMPI clinical scales. Several have significant negative correlations, and these same variables are positively associated with *Es*. *W*, *d*, *M*, *k*, *Fc*, *CF*, and *P* average about .40 with this variable. It will be noted that *Es* has a respectable correlation with Verbal IQ, a finding which supports Barron's observations. As had been demonstrated in earlier studies, the MMPI scales are markedly intercorrelated.

The intercorrelation matrix was analyzed by Thurstone's centroid method, and the analysis was continued until five factors had been extracted, at which time the residuals varied from —.18 to +.20.[3]

The resulting factorial patterns were rotated orthogonally and obliquely, an oblique solution being obtained by the trial vector method [3, p. 204]. Unrotated and rotated factors (oblique and orthogonal) are shown in Tables 2, 3, and 4. Descriptions are based on consideration of all three solutions.

Factor I contains high loadings on most of the Rorschach determinants and does not include MMPI scales. It can best be considered

Table 3

Orthogonal Factors

Variable No.	Symbol	I′	II′	III′	IV′
1	*W*	46	70	−10	09
2	*D*	90	−15	−06	−02
3	*d*	66	07	−28	−06
4	*M*	61	13	−26	00
5	*FM*	45	−02	−33	−03
6	*m*	40	16	−47	20
7	*k*	53	33	−03	−01
8	*K*	23	55	−16	27
9	*F*	83	−17	20	−13
10	*Fc*	28	32	−17	12
11	*c*	56	29	01	−01
12	*C'*	23	30	−09	07
13	*FC*	51	14	−65	14
14	*CF*	40	59	00	−07
15	*C*	12	49	−20	−10
16	*P*	01	−02	−04	−03
17	*R*	100	04	−27	−01
18	*VbIQ*	37	40	00	−03
19	*L*	19	02	42	−11
20	*F*	05	−67	−25	53
21	*K*	28	38	39	04
22	*Hs*	21	−52	07	63
23	*D*	14	−49	03	64
24	*Hy*	20	−29	48	53
25	*Pd*	−10	−08	03	86
26	*Pa*	−10	−33	−07	70
27	*Pt*	08	−69	−14	67
28	*Sc*	−01	−58	−22	74
29	*Ma*	−08	−05	−59	39
30	*Es*	28	77	09	−45
31	*A*	−02	−40	−09	83
32	*R*	−16	11	51	16

[2]Since the sample used was primarily composed of cases selected for a previous analysis of Rorschach determinants [10], the original correlations between Rorschach variables were utilized in the present matrix.

[3]The loadings on Factor V were so low that it was not included in later matrices.

Table 4

Oblique Factors

Variable No.	Symbol	I″	II″	III″	IV″
1	*W*	33	82	−12	−23
2	*D*	91	36	−48	04
3	*d*	65	41	−35	−10
4	*M*	60	42	−36	−06
5	*FM*	48	21	−39	−01
6	*m*	37	30	−49	11
7	*k*	43	60	08	−15
8	*K*	10	59	07	01
9	*F*	86	33	−37	−05
10	*Fc*	21	41	−12	−03
11	*c*	51	55	−17	−13
12	*C′*	18	32	−26	−06
13	*FC*	50	36	−48	06
14	*CF*	27	72	12	−18
15	*C*	02	46	06	−11
16	*P*	06	−06	−35	−03
17	*R*	99	60	−44	−02
18	*VbIQ*	27	54	08	−18
19	*L*	16	15	20	−11
20	*F*	11	−47	−12	76
21	*K*	16	51	33	−12
22	*Hs*	19	−19	25	77
23	*D*	09	−18	40	78
24	*Hy*	11	02	54	60
25	*Pd*	−18	−05	09	82
26	*Pa*	−11	−27	−04	77
27	*Pt*	13	−45	−09	89
28	*Sc*	01	−45	−19	90
29	*Ma*	−05	−11	−39	37
30	*Es*	19	71	−09	−73
31	*A*	−05	−24	17	91
32	*R*	−17	09	55	10

a general Rorschach productivity factor. A similar factor was found in previous studies, but in the course of rotation much of its variance was spread among other factors. As can be seen, rotation did not appreciably alter the factor in the present study. This dimension accounts for a high percentage of the total variance in the intercorrelation matrix, and its presence indicates that other factors must be treated conservatively.

Factors II and III are bipolar factors indicating that relationships between the two tests as they are now scored are probably inverse. That is, in Factor II, many of the MMPI scales have high negative loadings, particularly *Pt* and *Sc*, while the factor has positive loadings on Rorschach *W*, *CF*, and the *K* variables, Verbal IQ, and MMPI *Es* and *K*. The

Table 5

Correlations between Oblique Factors

Factor	I″	II″	III″	IV″
I″	1.00			
II″	38	1.00		
III″	−42	15	1.00	
IV″	−05	−26	−18	1.00

oblique and orthogonal positions are quite divergent. The peak loadings on the two solutions are similar, but the oblique axis includes a wider range of Rorschach determinants and greater productivity. As shown in Table 5, the latter solution results in correlation between Factors I and II.

From the loadings on Factor II, it seems reasonable to postulate that patients who are productive, giving a wide range of Rorschach determinants, who can organize the blots into wholes, and who use color with some freedom will tend to deny concern with themselves, dissatisfaction, pessimism, and worry, thus achieving a high *Es* score on the MMPI. In this case, Factor II may be tentatively labeled "ego strength."

Factor III is also bipolar, including significant loadings on MMPI *R* (repression), *L*, *K*, *D*, and *Hy*, along with Rorschach *F*. Its opposite pole is made up of *Ma* and Rorschach *R*, *D*, *M*, *FM*, *m*, and *FC*. Perhaps it may best be labeled "expressive-repressive."

Factor IV, a general MMPI dimension, is closely related to the first factor obtained in the several factorial studies mentioned above. It seems to involve a set to say psychiatrically negative things about oneself.

Summary and Conclusions

1. MMPI, Rorschach, and Wechsler-Bellevue data on 100 psychiatric patients were selected at random from the files of the Clinical Psychology Section at Letterman Army Hospital.

2. The intercorrelation matrix was factor analyzed by the centroid method.

3. Four factors were obtained and described. While relationship between the two tests was not remarkable, there is evidence for an inverse one. Two factors demonstrated overlap between the two tests. Factor II, tentatively

called "ego strength," included high loadings on the MMPI "ego-strength scale," and a fairly wide range of Rorschach determinants, with negative loadings on the MMPI "complaint" scales. Factor III contained high loadings on the MMPI "repression" scale and Rorschach *F*, with negative loadings on the movement variables, *FC* and the MMPI *Ma* scale.

Factors I and IV were identified respectively with Rorschach productivity and the MMPI "general maladjustment" factor. They indicate that a considerable amount of variance in each test is not accounted for by covariation between them.

4. The data suggest that a Rorschach variable might be derived that would function as a reliable predictor of ego strength.

5. The MMPI findings are in agreement with results of previous factorial studies.

Received August 17, 1953.

References

1. Altus, W. D. Some correlates of the group Rorschach and the schizophrenic scale of the group MMPI among two groups of "normal" college students. *J. consult. Psychol.*, 1948, **12**, 375–378.
2. Barron, F. An MMPI scale which predicts response to psychotherapy. *J. consult. Psychol.*, 1953, **17**, 327–333.
3. Cattell, R. B. *Factor analysis.* New York: Harper, 1952.
4. Clark, J. Some MMPI correlates of color responses in the group Rorschach. *J. consult. Psychol.*, 1948, **12**, 384–386.
5. Ellis, A. Recent research with personality inventories. *J. consult. Psychol.*, 1953, **17**, 45–49.
6. Thompson, G. M. MMPI correlates of certain movement responses in the group Rorschachs of two college samples. *J. consult. Psychol.*, 1948, **12**, 379–383.
7. Tyler, F. T. A factorial analysis of fifteen MMPI scales. *J. consult. Psychol.*, 1951, **15**, 451–456.
8. Welsh, G. S. A factor study of the MMPI using scales with item overlap eliminated. *Amer. Psychologist*, 1952, **7**, 341-342. (Abstract)
9. Wheeler, W. M., Little, K. B., & Lehner, G. F. J. The internal structure of the MMPI. *J. consult. Psychol.*, 1951, **15**, 134–142.
10. Williams, H. L., & Lawrence, J. F. Further investigation of Rorschach determinants subjected to factor analysis. *J. consult. Psychol.*, 1953, **17**, 261–264.
11. Wittenborn, J. R. A factor analysis of Rorschach scoring categories. *J. consult. Psychol.*, 1950, **14**, 261–267.
12. Wittenborn, J. R. Level of mental health as a factor in the implications of Rorschach scores. *J. consult. Psychol.*, 1950, **14**, 469–472.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

The Diagnosis of Subclinical Schizophrenia[1]

Donald R. Peterson
University of Illinois

One way to refine a scheme of psychiatric classification is to examine patients who, when certain developments in the course of their illness have been considered, turn out to have been incorrectly diagnosed. Careful study of behavior recorded at the time of their diagnoses should permit more definite limitation of class boundaries and more accurate placement of other patients for whom a diagnosis is required.

This study is concerned with the behavior of a group of veterans who should have been diagnosed as "latent," "incipient," or "subclinical" schizophrenics, but who in fact received some other designation. The importance of making the classification correctly for the protection of other members of society, for promoting maximal economy in treatment, and generally for facilitating wise patient disposition should be obvious.

Subjects and Procedures

Group 1. A set of false negatives was obtained by searching the complete file of case records at a VA mental hygiene clinic for all male, white patients who met the following criteria:

a. They were examined by a psychologist and were seen for at least two interviews by staff members at the clinic.

b. They received some diagnosis which contained no derivative of the word "schizophrenia."

c. They were later sent to a psychiatric hospital and there were said to be schizophrenic.

Group 2. A sample of true negatives was selected by examining data previously recorded [4] for a sample of nonhospitalized patients at the same clinic. Each false negative in Group 1 was considered in turn, and was compared consecutively with each patient listed in the data for the nonhospitalized sample until one was found who had received a primary diagnosis identical with that given his counterpart in Group 1. For these cases, neither diagnosis nor outpatient status was subsequently changed. The sample was again limited to male, white veterans.

Group 3. An exhaustive sample of true positives was found by searching the clinic files and selecting those patients who, as in the other groups, had received psychological and psychiatric examination at the clinic, whose diagnoses were "latent," "incipient," or "subclinical" schizophrenia, or "schizophrenia in remission," and who later were sent to the hospital and received some variety of the diagnosis "schizophrenia" there. Once more, all cases were male and all were white.

Groups 1 and 2 each contained 33 cases. For Group 3, $N = 27$. Data were gathered with regard to age, education, Wechsler-Bellevue IQ, and Minnesota Multiphasic Personality Inventory (MMPI) profile, and the false negatives compared with each of the two control samples by means, chiefly, of a χ^2 analysis, as specified by McNemar [2, formula 87] for 2 by k tables. This method was employed for testing differences in education and for investigating the discriminatory effectiveness of six signs said by Meehl [3] to be commonly employed in differentiating MMPI psychotic profiles from neurotic ones. The signs, modified only to make objective counting possible, are as follows:

1. T scores on four or more of the clinical scales over 70
2. $F > 65$
3. $Sc > Pt$

[1] Published with the permission of the Chief Medical Director, Department of Medicine and Surgery, Veterans Administration, who assumes no responsibility for the opinions expressed or the conclusions drawn by the author.

4. *Pa* or *Ma* > 70
5. *Pa* or *Sc* or *Ma* > *Hs* and *D* and *Hy*
6. *D* > *Hs* and *Hy*

In addition, the *t* test was applied for all continuous variables [1, formula 5.03]. Where an *F* ratio suggested dissimilarity in population variance, a method for approximating *t* was used which involves no assumption of homogeneous variance [1, pp. 74-75].

Results and Discussion

The diagnoses actually received by the false negative group are presented in Table 1.[2] They seem to have been most often confused with cases of anxiety reaction or mixed psychoneurosis.

Table 1

Primary Diagnoses of Patients in the False Negative Group

Diagnosis	Number
Psychoneurosis, mixed or unclassified	9
Anxiety reaction	9
Conversion reaction, hysteria or somatization reaction	4
Obsessive-compulsive reaction	3
Inadequate personality	1
Nonpsychiatric diagnoses	7

Differences in age, education, and intelligence were not significant either in the comparison between Groups 1 and 2 or that between Groups 1 and 3. Examination of *T* score means for the various MMPI scales suggested that the false negatives differ from the members of Group 2 on six of the scales, while they differ from the members of Group 3 with regard to three scales. Results of this type are not revealing enough to warrant interpretation.

Pattern analysis, even of the simplest kind, seems to yield far more meaningful information. A striking aspect of the profiles presented in Fig. 1 is the essential similarity between test patterns of the latent schizophrenics who were not correctly diagnosed and those who

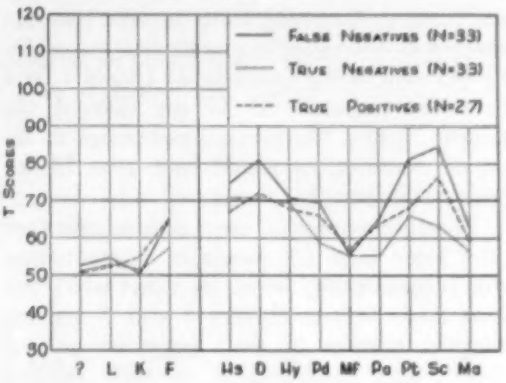

Fig. 1. Mean MMPI profiles.

were. When Meehl's psychotic signs were tested, not one produced a significant χ^2. In terms of the test pattern, however, the patients in Group 1 are obviously very much *unlike* the sample of true negatives, even though, as a group, they had been given identical diagnoses. Every one of the six signs yielded a χ^2 significant at or beyond the .05 level. In view of the informal way in which the criteria were established, the results are surprisingly good.

After-the-fact examination of the profiles leads me cautiously to suggest utility for one sign not included in this list or any other I know. It is the relationship between *Hy* and *Pd*—in inferential statement, a relationship between the repressive, inhibitory tendencies of the patient, and his propensity for "acting out" whatever impulses he may have. The "significance" which it yields here has little meaning, but the sign does not appear to be mentioned in other studies [4, 5] and might be worth including in future investigations of this type.

The implications furnished by the data are clear. Adequate information was there for the diagnosticians to use; it just was not employed as effectively as it might have been. By assigning to all cases one point for the presence of each sign, it is possible to derive a total "psychotic" score for any patient. When this was done, and a cutting score for class designation placed between one and two, it was

[2] A table of means, variance estimates, and results of significance tests for age, IQ, and the MMPI scales has been deposited with the American Documentation Institute. Order Document No. 4203 from the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., remitting in advance $1.25 for 35 mm. microfilm or $1.25 for 6 x 8 in. photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

found that 88 per cent of the incorrectly diagnosed patients in Group 1 would have been accurately placed. This increase, however, would have been gained at the expense of incorrect diagnosis for 39 per cent of the cases in Group 2. A cutting score between two and three would have led to correct designation for 67 per cent of the members of Group 1, and incorrect diagnosis for 18 per cent of the Group 2 cases. The over-all percentage of accurate classification would have been 74 for both cutting points.

These figures are far short of ideal, but they offer some hope for nosologic differentiation that is considerably better than that which has been made before. Groups 1 and 3 are exhaustive samples of their respective patient classes, and consideration of their relative sizes suggests that the diagnosis of subclinical schizophrenic is missed more often than it is correctly made.

That the problem with which we deal is one of inaccurate diagnosis, and not one of essential change in the patients themselves, is suggested by the results of follow-up scrutiny of 19 members of Group 1 who were retested at the hospital. Neither configural analysis of the type described above nor a trend analysis of the direction of sign movement produced statistically reliable results. Either the patients did not change much or the MMPI is too insensitive an instrument to measure any changes that did occur.

Summary

Refinement in psychiatric classification can be effected through specification of the behavioral characteristics of groups of false positives and false negatives. By isolating cases who are found, in the light of later developments, to have been incorrectly diagnosed, and by measurement of the behavior they exhibited before these facts were known, it is possible to improve diagnostic precision.

The present study is concerned with the diagnosis of subclinical schizophrenia, an important but difficult diagnosis for the clinician to make. Thirty-three patients had been examined at an outpatient clinic, and had received there some diagnosis containing no form of the word "schizophrenia." They were later hospitalized and were designated schizophrenic by staff members at the institution.

Comparisons were made between this group and another group of clinic patients, matched for initial diagnosis, with regard to age, education, Wechsler-Bellevue IQ, and performance on the MMPI. Similar comparisons were made between the criterion group and an exhaustive sample of 27 outpatients who, like the first group, were later hospitalized and diagnosed schizophrenic, but for whom the diagnosis had been correctly made at the clinic.

No differences in age, education, or intelligence were found. Neither did any clearly interpretable results emerge from analysis of the MMPI by means of the *t* test. Configural analysis through application of a group of psychotic signs, however, indicated that the MMPI patterns of the false negatives were essentially like those of the patients for whom the diagnosis, subclinical schizophrenia, had been accurately applied, and were substantially different from those produced by the group of patients with whom they had originally been classed. Diagnosis made in terms of these signs would have led to a considerable increase in classificatory precision.

Received August 14, 1953.

References

1. Johnson, P. O. *Statistical methods in research.* New York: Prentice-Hall, 1949.
2. McNemar, Q. *Psychological statistics.* New York: Wiley, 1949.
3. Meehl, P. E. Profile analysis of the MMPI in differential diagnosis. *J. appl. Psychol.,* 1946, **30,** 517–524.
4. Peterson, D. R. Predicting hospitalization of psychiatric outpatients. *J. abnorm. soc. Psychol.,* 1954, **49,** 260-265.
5. Wiener, D. N. A control factor in social adjustment. *J. abnorm. soc. Psychol.,* 1951, **46,** 3–8.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

The Relationship between Anxiety Level and Several Measures of Intelligence

Joseph D. Matarazzo, George A. Ulett, Samuel B. Guze, and George Saslow
Division of Psychosomatic Medicine and Departments of Medicine and Neuropsychiatry, Washington University School of Medicine

In an earlier study we reported the results of an experiment involving learning as a function of increasing levels of anxiety [4]. The study used 101 college sophomore *S*s, who were divided into seven groups, from low to high anxiety level, according to each individual's score on the Taylor Scale of Manifest Anxiety [8]. Individuals with Taylor raw scores between 1 and 5 were given an anxiety rating of 1, while those with raw scores of 6-10, 11-15, 16-20, 21-25, 26-30, 31-35 were given ratings of 2, 3, 4, 5, 6, and 7, respectively. There was one *S* with a score of 40, who was included with those given an anxiety rating of 7. Analysis of the Taylor raw score distribution yielded, reading from 1 to 7, the following distribution of *S*s along the seven-point scale: 4, 23, 27, 15, 17, 7, and 8. The results indicated that human maze learning had a significant curvilinear relationship to increasing levels of anxiety when total time to learn the maze was the criterion, and a significant rectilinear relationship when number of trials was the criterion of learning. Thus, when time was the measure of learning, *S*s in the middle anxiety range were superior to *S*s in both the high and low anxiety groups, whereas with the trials measure, *S*s at the high end were inferior to *S*s in the middle and low ends.

In an effort to evaluate whether differences in anxiety level were in fact responsible for the differences found in the mean learning abilities of the seven groups or whether the obtained differences were due to differences in intelligence levels of the groups, two available criteria of intelligence were used. Each *S* had been given the 1949 edition of the American Council on Education (ACE) Psychological Examination for College Freshmen, furnishing one criterion. The other criterion was each *S*'s grade-point average for the previous year, his freshman year.

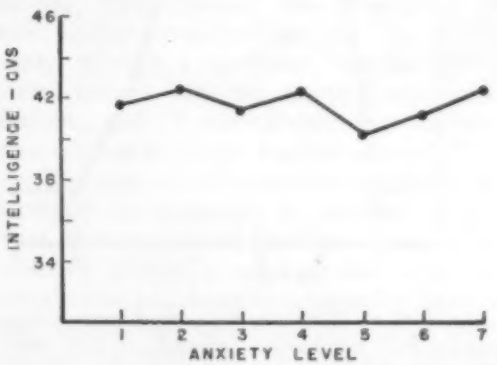

Fig. 1. Mean grade-point average for each of seven groups ranging from low anxiety (group 1) to high anxiety (group 7).

When the grade-point criterion was plotted as a function of anxiety level, no differences were found; i.e., grades had no significant relationship to level of anxiety. This can be seen in Fig. 1 where the mean grade-point averages for each of the seven groups are plotted. The function is essentially a straight line parallel to the abscissa. The Pearson *r* for these data, an insignificant value of —.08, confirmed the graphic impression. The value of the unbiased correlation ratio, epsilon, [5, pp. 319-324] was .00, and the chi-square test of linearity of regression [5, p. 319] based on this value and the corresponding Pearson *r* was .00, indicating that the function does not, in fact, deviate from a straight line. Analysis of variance yielded an insignificant *F* value of .88. The use of *t* tests based on the pooled variance revealed only one significant difference, groups 4 and 5 at the .05 level, and since 21 tests were performed this can be considered a chance

occurrence. Thus the conclusion of no relationship between grades and anxiety level appears warranted.

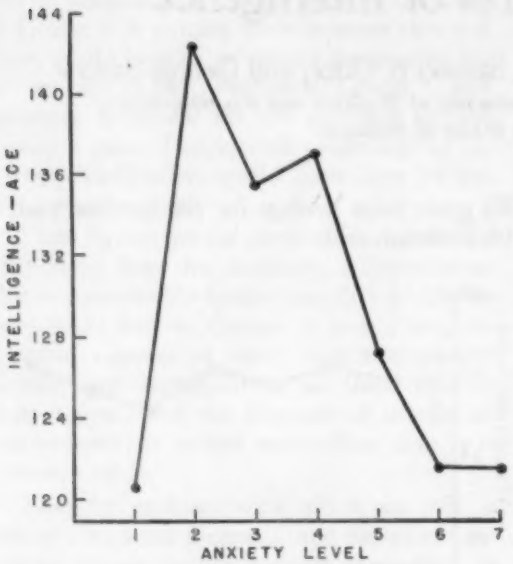

Fig. 2. Mean ACE score for each of seven groups ranging from low anxiety (group 1) to high anxiety (group 7).

The second criterion of intelligence, ACE scores, was analyzed next, as shown in Fig. 2. The apparent curvilinear relationship between total ACE score and increasing level of anxiety is strikingly similar to the actual function obtained for the *time* measure of learning and anxiety level in the earlier study [4]. However, unlike the time measure data of the earlier study, extensive statistical analysis revealed that the relationship between ACE score and anxiety level was in fact a rectilinear and not curvilinear one. Thus, the Pearson r between ACE score and anxiety level was —.25, a value significant at the .01 level of confidence, indicating a slight but nevertheless significant negative relationship between ACE score and anxiety level. The value of epsilon of .25 (.01 level of confidence) is equal to the value of r, and therefore the value of the chi-square test of linearity is .00, indicating a rectilinear rather than a curvilinear relationship.

As was pointed out in our earlier study, analysis of variance is limited as a test of differences among means with data of this type, for the reasons pointed out by Webb and Lemmon [9, 10]. Thus, with the ACE data of Fig. 2, we obtained an F value of 1.90, which is short of the 2.17 value needed for significance at the .05 level. Despite this lack of significance for F, two-tailed t tests based on the pooled variance yielded four significant differences: groups 1 and 2 at the .05 level and groups 2 and 5, 2 and 6, and 2 and 7 each at the .02 level of confidence. These values indicate that we are dealing with true differences rather than chance occurrences. However, the limitations of analysis of variance when one variable is ordered (our anxiety groups) should be kept in mind. Correlational techniques, like those reported above, are thus the methods of choice.

As with the data of our earlier study, the question of the small N (four Ss) in group 1 was next raised. It is obvious that the N's unduly influence the nature of the curve shown in Fig. 2, since the chi-square test results indicate that the relationship is a rectilinear one, i.e., that the essentially straight line from group 2 to 7 is probably a better indication of the relationship between ACE score and anxiety level. In an effort to test this, the data were regrouped as in the earlier study in an effort to increase the N in group 1. Since the 101 Ss had a range of Taylor raw scores from 1 to 40, it was decided that grouping them into five new groups in intervals of eight Taylor points (1–8, 9–16, 17–23, 24–32, 33–40) would be justifiable. The regrouping now placed 13 Ss in group 1, 47 in group 2, 23 in group 3, 14 in group 4, and 4 in group 5. Thus, while group 1 changed from 4 Ss to 13, group 5 was reduced to 4. Since it was felt that the N in group 1, and not those at the high end of the anxiety scale, was crucial, the 4 Ss in group 5 were included with the 14 Ss in group 4, thus yielding four new groups.

Figure 3 presents the relationship between ACE scores and these four differing levels of anxiety. As was predicted from the relationship between the r and epsilon obtained for the seven groups, the use of broader categories better represents the relationship between anxiety level and ACE score. Statistical analysis bears this out.

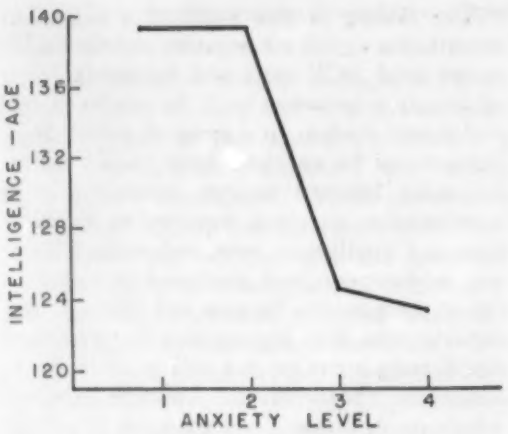

Fig. 3. Mean ACE score for each of four groups ranging from low anxiety (group 1) to high anxiety (group 4).

The value of r for the data of Fig. 3 was, of course, still —.25 (significant at the .01 level). The new value of epsilon was .28 (also significant at the .01 level). As the slight difference between these correlations would indicate, the value of the chi-square test of linearity of regression, 2.91, did not approach significance. Analysis of variance yielded an F of 4.03 (.01 level), and analysis by t tests indicated that groups 1 and 2 were *each* significantly different from groups 3 and 4: groups 1 and 3 and 1 and 4 at the .05 level; groups 2 and 3 and 2 and 4 at the .01 level of significance. These analyses and those of Fig. 2 indicate that the relationship between ACE score and anxiety level is rectilinear, and the curve in Fig. 3 is the best representation of this relationship. Therefore it can be concluded that anxiety appears to be moderately, although significantly, correlated with total ACE score and that the higher the anxiety level the lower, on the average, is the obtained score on this test. Analysis of the two parts of the ACE revealed that the relationship between the Q and L scores and anxiety was approximately the same. Both yielded negative correlations comparable to the one shown in Fig. 3 for the total score.

Since there was a lack of agreement between the relationship of anxiety level and our two original criteria of intelligence, an explanation was sought. The Pearson r between total ACE score and grades was .51, significant at the .01 level. As this value is very near the mode of all reported correlations between these two measures [7, p. 120], it was not believed that our sample was responsible for the lack of agreement between the relationship obtained for anxiety and each of our two criteria. Instead the answer was sought in the difference between our two criteria.

As the ACE is a timed test, we wondered if the speed factor could account for the negative correlation obtained with increasing levels of anxiety. This hypothesis was strengthened by two further considerations: (*a*) our earlier study revealed a significant relationship between anxiety level and the time measure of learning, and (*b*) clinical experience indicates that anxious *S*s do not function efficiently under time pressure. We therefore decided to investigate this hypothesis by retesting *S*s, this time with a nontimed intelligence test. If the obtained function is a straight line parallel to the abscissa, i.e., shows no relationship between this measure of intelligence and anxiety level, we can conclude that the hypothesis that the time factor of the ACE accounts for its negative correlation with anxiety level is reasonable.

The nontimed intelligence test used was the *CVS*, an abbreviated version of the Wechsler-Bellevue developed by Hunt and his associates [3]. None of the three subtests of this scale, Comprehension, Vocabulary, and Similarities, is timed, and *S*s were allowed, in a group administration, to take as much time as they required to complete the scale.

The results, when this criterion of intelligence (total *CVS* score) is plotted against increasing levels of anxiety, are given in Fig. 4. The data are shown as a function of all seven groups. As the figure demonstrates, and actual analysis revealed, the use of the broader categories (four groups) gave an essentially similar curve. Inspection of Fig. 4 indicates that the hypothesis that the time factor of the ACE may be responsible for the obtained negative correlation appears to be borne out. Statistical analysis confirms this impression, as the values

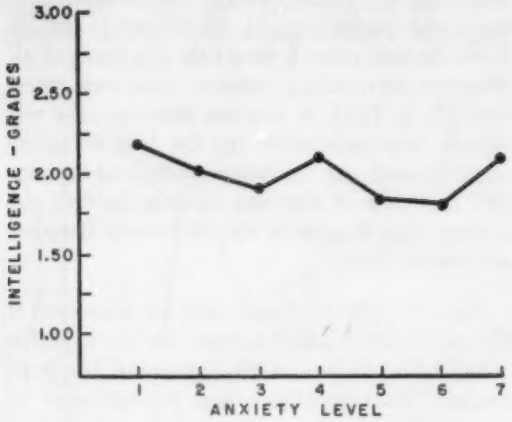

Fig. 4. Mean *CVS* score for each of seven groups ranging from low anxiety (group 1) to high anxiety (group 7).

of *r*, epsilon, and chi square[1] were —.07, —.14, and —1.72. None of these values approached significance. The value of *F*, .68, was insignificant, and *t*-test analyses revealed only one, and therefore chance, significant (.05 level) difference out of the 21 tests computed. Analysis of the relationship of *C*, *V*, and *S*, taken separately, as a function of anxiety level revealed three curves similar to Fig. 4.

It can therefore be concluded that intelligence, as measured by the *CVS*, like grades as a measure of intelligence, is not related to anxiety level. However, the relationship between *CVS* score and grades is not to be construed as a noncomplex affair, as the relatively low, although significant (.01 level) Pearson *r* of .34 between them demonstrates. This value is typical of those reported between measures of intelligence and grades [7, p. 145]. Finally, the Pearson *r* between ACE and *CVS* scores of .57 (.01 level) is, again, the typical finding [7, pp. 117-118]. The lack of even stronger relationships among these various measures indicates the complexities introduced by the interaction of personality and intellectual variables, as well as the limitations of validity and reliability inherent in the criteria themselves.

[1] The negative sign results from the chance negative value of epsilon. This is discussed by Peters and Van Voorhis [5, p. 355].

The finding in this study of a slight but nevertheless significant negative correlation between total ACE score and increasing levels of anxiety is consistent with the results of several recent studies. In a series of publications, Sarason and his associates have studied the relationship between anxiety, measured by a questionnaire sampling reactions to examinations and intelligence tests, and various learning, achievement, and psychometric tasks. In one of these studies Sarason and Mandler [6] demonstrated that low-anxious *S*s performed significantly better on two college entrance examinations (Mathematical Aptitude Test and Scholastic Aptitude Test) than did *S*s scoring at the high-anxious end. As in our study, a high level of anxiety was associated with lower performance on these two aptitude tests. Furthermore, these authors found no significant relationship between their anxiety measure and actual grade average for the previous year's work — a finding consistent with ours. And finally, their obtained significant correlation of .21 (.05 level) between anxiety level and a 20-minute timed intelligence test, the Henmon-Nelson Test of Mental Ability, college level, is again consistent with our correlation of .25 (.01 level) between anxiety level and the timed ACE scale. Thus, these two studies, completed independently of each other and using different indices of anxiety as well as different measures of intellectual functioning, yielded essentially similar results.

Two recent Rorschach studies report results consistent with these findings. In a controlled experiment, Beier showed that the effect of induced anxiety, a standardized Rorschach personality interpretation designed to promote stress, was to significantly decrease performance on such intellectual tasks as abstract reasoning, sorting, and mirror tracing [1]. It is of interest that several of Beier's measures were timed. In the other experiment, Gaier obtained an index of anxiety-readiness for each of his 11 *S*s on the basis of their Rorschach protocols and related this characteristic to each *S*'s performance on a variety of scholastic and other type tests [2]. He found a rank-order correlation between anxiety-readiness and the *L* score of the ACE of —.37, and —.39 between the *Q* score and anxiety-readi-

ness. These results are again similar to those of our study.

The implication of these findings, and those of the present study, is not to be drawn from the magnitude of the correlations between a personality variable ("anxiety") and various measures of "intelligence." These correlations are too low. Rather, the importance of the findings lies in the fact that they are consistent with clinical observation, and they point out an area of fruitful research once new and more valid measuring instruments are devised. The interaction of personality and intellectual variables, so-called, will receive much needed study once the problem of objective measures has been solved, especially for personality. The Taylor scale is one small step in that direction.

Summary

Three criteria of intelligence were each evaluated as a function of increasing levels of anxiety for 101 *Ss* who were divided into groups according to their scores on the Taylor Scale of Manifest Anxiety. The results indicated no relationship between anxiety level and the first and third of these criteria: grade-point average and score on the *CVS* scale. The relationship between the second criterion, score on the ACE scale, and increasing levels of anxiety was found to be a slight but nevertheless statistically significant (.01 level) negative one (r of —.25). The fact that the ACE is a timed scale appeared to be a reasonable hypothesis to account for the negative correlation obtained. The use of the third criterion, the untimed *CVS* scale, seemed to bear out this hypothesis.

Received September 14, 1953.

References

1. Beier, E. G. The effect of induced anxiety on flexibility of intellectual functioning. *Psychol. Monogr.*, 1951, **65**, No. 9 (Whole No. 326).
2. Gaier, E. L. Selected personality variables and the learning process. *Psychol. Monogr.*, 1952, **66**, No. 17 (Whole No. 349).
3. Hunt, W. A., French, Elizabeth G., Klebanoff, S. G., Mensh, I. N., & Williams, M. Further standardization of the CVS individual intelligence scale. *J. consult. Psychol.*, 1948, **12**, 355–359.
4. Matarazzo, J. D., Ulett, G. A., & Saslow, G. Human maze performance as a function of increasing levels of anxiety. *J. gen. Psychol.*, in press.
5. Peters, C. C., & Van Voorhis, W. R. *Statistical procedures and their mathematical bases.* New York: McGraw-Hill, 1940.
6. Sarason, S. B., & Mandler, G. Some correlates of test anxiety. *J. abnorm. soc. Psychol.*, 1952, **47**, 810–817.
7. Super, D. E. *Appraising vocational fitness.* New York: Harper, 1949.
8. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.
9. Webb, W. B., & Lemmon, V. W. A qualification in the use of analysis of variance. *Psychol. Bull.*, 1950, **47**, 130–136.
10. Webb, W. B., & Lemmon, V. W. A sequel to the notes of Patterson and Diamond. *Psychol. Bull.*, 1952, **49**, 155.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Comparison of Two Forms of the Manifest Anxiety Scale

Joyce B. McCreary and A. W. Bendig[1]
University of Pittsburgh

Taylor [1] reported the development of a Manifest Anxiety Scale (MAS) consisting of 50 items drawn from the MMPI. Scores are obtained either by scoring only these items from the full-length MMPI or by embedding them in 175 additional filler items and giving the resulting 225-item test. Taylor [1, p. 287] mentions that distributions of scores obtained by these two methods are not similar, and the correlation between scores from the two forms was only .68 ($N = 282$). Commonly, the 225-item test is used to identify anxious subjects, but this long test is not economical of either the subject's or the experimenter's time. Preferably only the 50 scored items should be given to *S*s. However, the results previously quoted suggest that scores from this shortened form might not be comparable to those from the longer forms.

Students ($N = 56$) were given two forms of the MAS. Form A contained the 225 items described above, but only 50 of them, scattered irregularly throughout the form, were scored. Form B contained only the 50 scored items from Form A, all other items being omitted. In one section ($N = 22$), Form A was given first followed immediately by Form B. In the other section ($N = 34$), the forms were administered in reverse order.

The 112 scores were converted to normalized *T* scores ($M = 50$, $SD = 10$), based upon the performance of the norm group of 1,971 students at the University of Iowa [1]. Analysis of variance showed that the variation between classes (which was confounded with the order of presentation of the two test forms) and between forms was not significant. Differences between the trials were significant at the .05 level of confidence. Comparison of trial means indicated that *S*s tended to show lower (less anxious) scores on the second test given them regardless of form. The correlation between Form A and Form B transformed scores was .95: a similar correlation between pairs of raw (untransformed) scores was .96.

The shortened form of the MAS (Form B) gave results almost identical with those obtained from the long form (Form A). Evidently either form can be used for the selection of anxious *S*s with equivalent results. Since Form B is more parsimonious in terms of class time, its use is preferable. The significant between-trials term in our analysis of variance suggests an explanation of the differences found by Taylor between scores from the full-length MMPI and those from the 225-item MAS. Many of the MAS items are near the end of the MMPI, and *S* has had to answer several hundred other items before reaching them. The "fatigue" effect found in our study may be accentuated and compounded when an even longer test is used.

Brief Report
Received February 8, 1954.

[1]An extended report of this study may be obtained without charge from A. W. Bendig, Department of Psychology, University of Pittsburgh, Pittsburgh 13, Pa., or for a fee from the American Documentation Institute. To obtain it from the latter source, order Document No. 4211 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., remitting in advance $1.25 for 35 mm. microfilm or $1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

Reference

1. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285–290.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Rater Experience and the Reliability of Case History Ratings of Adjustment

A. W. Bendig and Janine Sprague
University of Pittsburgh

In Guilford's discussion of the optimum length of rating scales [5, pp. 267-268], the hypothesis is advanced that longer scales, with seven to nine categories, should be used only with experienced and well-motivated judges, whereas shorter scales should be used with relatively naive judges. Presumably, longer scales place too great a burden on the discriminatory ability of inexperienced raters. Although widely adopted as a guiding principle in the construction of rating scales, this rational hypothesis has, as far as we have been able to determine, never been subjected to experimental test. Previous research [2, 3] has confirmed several of Guilford's strictures on constructing rating scales and has questioned others. The present research was designed to clarify the relationships between scale reliability, the past experience of the rater, and length of the scale. Rater experience is used here as task-relevant experience, i.e., the amount of past exposure to and familiarity with the kind of judging task required of the rater.

Procedure

Stimuli. Ten short (approximately 100-word) case histories were abstracted from several commonly used textbooks in abnormal psychology. The cases ranged from mildly disturbed neurotic to institutionalized psychotic. Most of the cases were reduced in length, with the more technical wording either eliminated or simplified. The cases were mimeographed as a booklet with a face sheet of instructions to the raters. The instructions requested the rater (*a*) not to discuss the cases with his friends before rating, (*b*) to make his ratings as objective as possible, and (*c*) to make judgments independent of previously rated case histories.

Scales. Numerical rating scales with 3, 5, 7, or 9 categories were mimeographed on separate sheets with spaces for recording the ratings of the case histories. Three simple verbal anchors were used, one under each of the two end scale categories and one under the center. The anchors were:

Slightly maladjusted
Moderately maladjusted
Extremely maladjusted

Subjects. The total group of 208 raters was composed of four experience groups: 108 students in introductory psychology; 62 in the second semester of abnormal psychology; 29 graduate students in psychology; and 9 psychologists having the Ph.D. degree and several years experience in clinical or counseling work. Most of the introductory students were freshmen or sophomores, and they rated the cases sometime after the period in the course devoted to personality and behavior disorders. Almost all of the abnormal psychology students were juniors and seniors, with a large minority majoring in psychology. The graduate students ranged from first-year to fourth-year students, with a strong majority in clinical and counseling areas. All of the Ph.D. raters were affiliated with this university or some neighboring agency in a clinical or counseling capacity and had a minimum of three years' experience in service activities. For the first three groups, the case histories and rating sheets were passed out in class, and students were permitted to rate outside of the classroom. The Ph.D. raters were contacted individually and also rated at their leisure. Scales with 3, 5, 7, and 9 categories were randomly distributed to the introductory and abnormal psychology students. Because of the smaller numbers of raters, and 5- and 7-category scales were randomly distributed to the graduate students, and only the 7-category scale was given to the Ph.D.'s.

Table 1

Reliability Analysis of Adjustment Ratings

Rater group	*N*	Number of categories	Rater reliability			Rater bias	
			F	Group *R*	Individual *r*	*F*	*R*
Introductory	27	3	15.64**	.94	.35	.76	—.31
psychology	26	5	17.76**	.94	.39	1.15	.13
students	21	7	19.68**	.95	.47	2.17**	.54
	34	9	39.97**	.97	.53	2.69**	.63
Abnormal	13	3	15.20**	.93	.52	3.44**	.71
psychology	15	5	12.99**	.92	.44	2.76**	.64
students	18	7	29.83**	.97	.62	4.67**	.79
	16	9	16.60**	.94	.49	4.79**	.79
Graduate	13	5	18.76**	.95	.58	2.25*	.56
students	16	7	23.94**	.96	.59	2.32**	.57
Clinical Ph.D.	9	7	17.65**	.94	.65	2.07*	.52

* Significant at the .05 level.
** Significant at the .01 level.

Results

The reliability of the individual raters for each of the 11 subgroups was estimated by analysis of variance techniques [2, 3]. These results can be found in Table 1.

In general, individual rater reliability increased with higher levels of education, with the Ph.D. group showing the largest reliability ($r = .65$). Consideration of the interaction of rater experience and length of scale discloses an interesting phenomenon. Introductory students show a smooth linear increase in reliability as the scale lengthens until, with the 9-category scale, they surpass the abnormal psychology students using a 9-category scale and almost reach the reliability level of the graduate students. The four subgroups of abnormal psychology raters fluctuate in reliability, showing their highest reliability with the 7-category scale and their lowest with 5 categories. This greater fluctuation may be due, in part, to the smaller numbers of *S*s per group. However, the general reliability trend is upward for longer scales, but the rate of increase (slope of the best-fitting straight line) is not as steep for the abnormal as it is for the introductory psychology *S*s. There was no difference in the rater reliability found for the two graduate student subgroups using 5- and 7-category scales, and both groups showed slightly smaller reliability than did the Ph.D. group.

As to the magnitude of constant errors in the ratings of the subgroups (rater bias), the introductory psychology *S*s show a progressive increase with longer scales, surpassing the Ph.D. group in bias with a 7-category scale, but not surpassing the graduate student subgroups until the 9-category scale was reached. The abnormal psychology raters consistently showed a greater bias for every scale length, but again showed some fluctuation between these subgroups. Of all four subgroups using the 7-category scale, the Ph.D. raters showed the smallest bias.

The mean ratings of the ten case histories by the 11 subgroups were intercorrelated and the matrix of coefficients found.[1] All of the 55 correlations were extremely high (the median being .90), with 54 significant at the .01 level, $df = 8$, and the remaining correlation, $r = .73$, significant at the .05 level. A centroid factor analysis was performed on these correlations, replicating the analysis twice to stabilize the communalities; the unrotated factor loadings can be found in Table 2. Factor extraction was terminated after two factors because of small residual correlations. Factor I was a broad general factor accounting for an average of 91 per cent of the variation in mean

[1] A one-page table showing intercorrelations between mean adjustment ratings of eleven groups of raters has been deposited with the American Documentation Institute. Order Document No. 4206 from the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., remitting in advance $1.25 for 35 mm. microfilm or $1.25 for 6 x 8 in. photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

Table 2

Factor Loadings from Analysis of Intercorrelations between Mean Adjustment Rating of Rater Groups

Group	Number of categories	Factor* I	Factor* II	h^2
Introductory psychology students	3	94	—37	102
	5	95	—24	96
	7	97	—18	97
	9	95	—23	96
Abnormal psychology students	3	90	15	83
	5	95	17	93
	7	96	20	96
	9	97	15	96
Graduate students	5	95	26	97
	7	97	22	99
Clinical Ph.D.	7	97	18	97
Per cent of total rater variance		91	5	96

* Decimal points omitted.

case history ratings and appeared to be a function of the length of the scale. Subgroups using a 7-category scale consistently showed higher loadings on Factor I, and the two subgroups using 3-category scales had the smallest loadings. Factor II was associated with 5 per cent of the remaining intragroup rating variance and was clearly related to the experience background of the raters. This "experience" factor was bipolar, with the four introductory subgroups having negative loadings and the other subgroups spaced along the positive side of the vector. In all, these two factors accounted for 96 per cent of the variance between mean case history ratings.

The nonparametric rank test of Kruskal and Wallis [6] was used to test the above interpretation of I as a scale-length factor and II as an experience factor. The 11 subgroups were ranked according to their loadings on each of the factors, and the sums of ranks of the four introductory, four abnormal, and three graduate student-Ph.D. subgroups on Factor I and on Factor II were found. Applying the rank test to the sums of ranks of the three experience groups on Factor I gave an H value of 1.30, which is not significant at the .05 level, while applying the same test to the three sums of ranks on Factor II yielded an H of 8.91, which is significant at the .01 level [6, p. 615]. Again the 11 subgroups were ranked on each factor and the sums of ranks for each of the four lengths of scale obtained. The H test for Factor I was 7.48 (significant at the .05 level) and for Factor II, H equalled 3.17 (not significant at the .05 level). These tests of significance support our interpretation of Factor I as related to scale length and Factor II as associated with rater experience. Objection may be raised to our interpretation of Factor II since our experience groups did show a general increase in rater reliability from introductory students to Ph.D. clinicians. The interpretation of Factor II as a "reliability" factor does not appear plausible when we note the large overlap in rater reliability between the introductory and abnormal student groups (Table 1) and the complete lack of overlap between these groups on Factor II.

Finally, the above Factor II results suggested the hypothesis that the introductory-abnormal-graduate student groups should show a progressive increase in the correlation of their mean case history ratings with those of the Ph.D. group. The average correlations of the two introductory, two abnormal, and two graduate student groups who used the 5- and 7-category scales with the Ph.D. group were found and tested for significant differences by the r-to-z technique described by Edwards [4, pp. 132-136]. The critical ratios resulting from these comparisons can be found in Table 3. The introductory group showed a significantly smaller average correlation with the Ph.D. group than did the abnormal or graduate student groups, on a one-tailed test corresponding to the above directional hypothesis. However, the comparison of average correlations of the abnormal and graduate student groups with the Ph.D. group showed no significant difference.

Table 3

Significance of the Differences between Average Correlations of Three Rater Groups with the Clinical Ph.D. Raters

Group	N	$\bar{r}$	Critical ratios: Abnormal	Critical ratios: Graduate
Introductory	47	.87	1.88*	3.37**
Abnormal	33	.96		1.07
Graduate	29	.97		

* Significant at the .05 point.
** Significant at the .01 point.

Discussion

The results of this study appear sufficiently unequivocal to question seriously the Guilford-sponsored statement that longer scales should be used only with experienced raters. Our results suggest just the opposite: for more reliable ratings longer scales should be used with inexperienced raters and (tentatively) short scales can only be resorted to when the raters have had experience in the type of judgment required. As a rationale to explain our results we suggest that inexperienced raters are more likely to disagree as to the scale placement of a particular stimulus than are experienced raters and disagree to the extent, approximately, of one scale category. With long scales this one-category disagreement represents a much smaller percentage of the total rating variance than with short scales. Since rater reliability is a function of the ratio of rater disagreement variance to the total variance, we would expect shorter scales, with smaller total variance, to yield less reliable ratings. However, experienced raters can better tolerate the compression of the scale and better maintain agreement with short scales.

Parenthetically it might be noted that our results are a tribute to the rating ability of the Ph.D. clinicians. Since we wanted to compare ratings from all levels of the experience continuum, both our case histories and scale anchors had to be designed to be comprehensible to the lowest common denominator — the introductory psychology student. Consequently, many of the Ph.D. clinicians justifiably felt that our case histories were too short and incomplete for them to make adequate judgments and that the points on our scales were inadequately defined. In spite of these limitations the average intercorrelation among the nine Ph.D. clinicians was .65, and their group judgment had a reliability of .94. Doubtless with more complete case histories and with a better definition of what we meant by "adjustment," these reliabilities would have been increased.

The factor analysis of the intercorrelations of mean ratings indicates that the largest influence in these correlations is the length of the scale. Factor I, which is fairly clearly related to scale length, accounts for 91 per cent of the total variability in the mean ratings. Factor II was even more clearly an experience factor, but only tapped 5 per cent of the rating variance. We must admit at this point that the difference between our major experience groups, which is related to Factor II, is composed of a number of different variables. Presumably there are differences in age, intelligence, educational training, motivation, and sex percentages between introductory, abnormal, graduate student, and Ph.D. groups. Which of these organismic variables are most responsible for differences in rating can only be a matter of speculation at this point. However, the academic selection processes that operate to winnow from introductory psychology a small group who will receive the Ph.D. also operate to increase the agreement among the raters as to what constitutes maladjusted behavior. This selective process appears to be most influential between the introductory and abnormal psychology courses and is probably related to whatever factors determine that students will become majors in psychology.

The increasing reliability of naive raters with longer scales found in this study contrasts markedly with the results of previous studies [2, 3]. All three of these studies required self-rating on the part of the rater, i.e., using the scale to report the rater's perceptions of his own feelings, preferences, etc. In the former studies, scale reliability did not markedly increase with longer scales. In the present research, the raters were asked to evaluate stimuli that were more objective and more independent of the rater, i.e., verbal descriptions of the behavior of other people. Presumably these ratings could still be affected by a projective mechanism, but it seems likely that somewhat less ego-involving stimuli would encourage greater interjudge agreement [1]. We would predict that naive or inexperienced judges would show a steeper gradient of increasing rater reliability with lengthened scales if even more objective stimuli were rated.

Summary

Raters ($N = 208$) judged the adjustment level shown in each of ten short case histories. Four rater groups were used: introductory psychology, abnormal psychology, psychology graduate students, and a group of post-Ph.D.

clinical psychologists. Rating scales with 3, 5, 7, or 9 categories were used. Results indicated a general increase in reliability with higher educational levels. Rater reliability increased sharply with longer scales for the introductory group and increased at a slower rate for the abnormal psychology group. Factor analysis of the intercorrelations among the groups indicated that a factor related to the length of the scale accounted for 91 per cent of group variance, and a second factor reflecting experience differences measured 5 per cent of the variance.

Received August 17, 1953.

References

1. Bendig, A. W. Inter-judge vs. intra-judge reliability in the order-of-merit method. *Amer. J. Psychol.*, 1952, **65**, 84–88.
2. Bendig, A. W. The reliability of self-ratings as a function of the amount of verbal anchoring and of the number of categories on the scale. *J. appl. Psychol.*, 1953, **37**, 38–41.
3. Bendig, A. W. Reliability and the number of rating scale categories. *J. appl. Psychol.*, 1954, **38**, 38–40.
4. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
5. Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.
6. Kruskal, W. H., & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Amer. statist. Ass.*, 1952, **47**, 583–621.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Personality Correlates of Undergraduate Field of Specialization[1]

Richard C. Teevan
University of Michigan

This paper attempts to investigate whether or not the personality of an individual is correlated significantly with his choice of major field in college. In recent years there have been two research programs which stimulated the present work: the survey of scientists by Knapp and Goodrich [4] and the investigations of eminent scientists by Roe [5, 6]. These studies attempted to explore the personalities and backgrounds of men who are presently working in the sciences.

Since the above investigations, among others, have shown that there is a correlation between personality and vocation, the question arises whether this correlation can be demonstrated before the vocation is actually entered or whether the vocation itself fits the personalities of its members into a common mold. The present paper throws light on this matter by using college students who have already expressed their preference for a major field but who have not been in that field long enough for it to affect their personalities. The fact that choice of a major in college is not necessarily followed by a similar vocation is a drawback, but one that might serve only to attenuate any obtained relationships.

Procedure

The diagnostic instrument in this study was the Blacky Pictures [2], a modified projective technique measuring psychosexual development. Eighty-five male students from Wesleyan University, ranging in age from 17 to 25 years, were tested. The subjects were chosen by their major field, with approximately the same number from each of the academic divisions. The major fields were broken down into three divisions: 1. music, art, German, French, and English ($N = 29$); 2. government, economics, and history ($N = 28$); and 3. geology, astronomy, biology, physics, and chemistry ($N = 28$). An attempt was made to obtain subjects who were sincerely interested in their field of endeavor, so that an average of B— was taken as the criterion.

The Blacky Pictures were administered in three sections, each section, insofar as possible, containing equal numbers of subjects from each division. There were two deviations from standard procedure: (*a*) Up to 15 seconds longer than the standard two minutes were allowed on the spontaneous stories; (*b*) the inquiry questions, rather than being flashed on the screen one at a time with a time limit for each question, were printed at the bottom of each page. Subjects were instructed not to glance at them until told to. An observational check and questioning at the end of the session showed good observance of these instructions.

Results and Discussion

The findings of this study are summarized in Tables 1 and 2. Table 1 gives the significance of the differences between divisions when Blacky disturbance scores are related to choice of major field. In Table 2 the percentages of each group showing disturbed and neutral responses are given for those Blacky dimensions on which significant differences were obtained.

The first division (English, art, languages) was found to be the highest of the three divisions on Oral Eroticism. Since the great majority of the subjects tested (22 out of 29) were from the English department, it is perhaps possible to treat them as literary persons. Bergler [1] observed 36 writers in psychoanalytic treatment and concluded that all writers

[1]This paper is abstracted from an honors thesis done at Wesleyan University under the guidance of Professor Robert H. Knapp.

Table 1

The Relationship between Blacky Scores of Disturbance and Choice of Major Field

Blacky category	p	Ranking of division by degree of disturbance		
		Most	Middle	Least
Oral Eroticism	.01	1	2	3
Oral Sadism	.05	2	1	3
Anal Expulsiveness	NS			
Anal Retentiveness	NS			
Oedipal Intensity	.06	2	1	3
Masturbation Guilt	NS			
Castration Anxiety	NS			
Identification Process	NS			
Sibling Rivalry	NS			
Guilt Feelings	.01	2	1	3
Ego Ideal	NS			
Narcissistic Love Object	NS			
Anaclitic Love Object	.07	2	1=3	1=3

Note.—Division 1—music, art, German, French, English; Division 2—government, economics, history; Division 3—geology, astronomy, biology, physics, chemistry.

Table 2

Percentage of Each Division Showing Disturbed (+) and Neutral (O) Responses to Blacky Dimensions

Blacky dimension	Division		
	1 Lit.	2 Soc.	3 Sci.
Oral Eroticism			
Disturbed (+)	61	31	18
Neutral (O)	39	69	82
Oedipal Intensity			
Disturbed (+)	57	66	36
Neutral (O)	43	34	64
Oral Sadism			
Disturbed (+)	46	55	18
Neutral (O)	54	45	62
Guilt Feelings			
Disturbed (+)	64	86	39
Neutral (O)	36	14	61
Anaclitic Love Object			
Disturbed (+)	11	31	11
Neutral (O)	89	69	89

are psychic masochists and therefore fixated at the oral erotic level. Boyd [3], using the Blacky Pictures on children, found more disturbance at the oral erotic level in good readers than in retarded readers. Thus the present findings from the Blacky Pictures are consistent with previous descriptions of literary people as passive and dependent.

The second division (economics, government, and history) was highest on the other four categories that showed differences: Oral Sadism, Oedipal Intensity, Guilt Feelings, and Anaclitic Love Object. Three of these categories contain a common element — emphasis on the mother. Thus Oral Sadism is related to aggression toward the mother (see Blacky Picture No. 2); Oedipal Intensity indicates an unresolved conflict over desire for the mother; and Anaclitic Love Object reflects a desire for the mother as a love object. Guilt Feelings, the fourth dimension on which the second division is highest, is based on the introjection of parental standards and also conceivably fits into this pattern. Thus the disturbances prevalent in Division 2 tend to center around the boy's relationship to his mother. Any interpretation of this set of findings is sufficiently speculative to allow the reader the freedom of his own projections in this regard. The facts themselves, however, seem quite clear.

Another striking point is the fact that Division 3 (the sciences) has the lowest disturbance scores on most of the Blacky categories. Since no artifacts can be uncovered to account for this finding, we can conclude that college students who select a major field within this division are less disturbed than those who major in the other fields used in this study.

Summary

This investigation was conducted to determine whether or not personality factors are correlated with choice of a major field in college. Three broad groupings of college majors were compared on scores derived from the Blacky Pictures [2]. The following significant differences were found:

1. The "literature" division had higher disturbance scores on Oral Eroticism than the other divisions.

2. The social sciences division had higher disturbance scores on Oral Sadism, Oedipal Intensity, Guilt Feelings, and Anaclitic Love Object.

3. The science division had the lowest disturbance scores on nearly all categories.

From the above results it is concluded that correlations between personality and vocation previously found by Roe [5, 6] for professional

groups can also be demonstrated during the period preceding entrance into a profession.

Received November 5, 1953.

References

1. Bergler, E. *The writer and psychoanalysis.* New York: Doubleday, 1950.
2. Blum, G. S. A study of the psychoanalytic theory of psychosexual development. *Genet. Psychol. Monogr.*, 1949, **39**, 3-99.
3. Boyd, R. D. Reading retardation as related to personality factors of children and their parents. Unpublished doctor's thesis, Univer. of Michigan, 1953.
4. Knapp, R. H., & Goodrich, H. *Origins of American scientists.* Chicago: Univer. of Chicago Press, 1952.
5. Roe, Anne. A Rorschach study of a group of scientists and technicians. *J. consult. Psychol.*, 1946, **10**, 317–327.
6. Roe, Anne. Psychological examinations of eminent biologists. *J. consult. Psychol.*, 1950, **13**, 225-246.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Congruence of Self and Ideal Self in Relation to Personality Adjustment

Thomas E. Hanlon, Peter R. Hofstaetter, and James P. O'Connor
The Catholic University of America

Recent studies of patient change during successful psychotherapy [4, 6, 7, 8, 9] have demonstrated an increase in the correlation between the client's self concept and his ideal self concept as measured by Roger's *Q* sort.[1] The hypothesis of the present study springs directly from this observation. If successful psychotherapy promotes better adjustment, a positive and significantly high correlation should exist between measures of adjustment and the congruence of the self concept and the ideal self. This correlation should also exist in a sample from a population in which departures from satisfactory adjustment are still within the "normal" range, that is, in a population whose members are not patients. Studies by Bills *et al.* [1] and by Draffen [2] suggest that the hypothesis will be confirmed.

Procedure

The subjects of this study were 78 male juniors in a Catholic high school in Wilmington, Delaware. The mean age of the group was 16.3 years, with a range from 14 to 18 years and a standard deviation of .78 years. The mean IQ, as measured by the Otis Self-Administering Tests of Mental Ability, Intermediate, Form B, was 104.3 with a range from 77 to 140 and a standard deviation of 11.4.

Since the *S*s were available for only short periods of time, measures of the self concept and ideal self concept were obtained by employing a modification of the rather lengthy *Q*-sort technique. The 100 self-referent statements were presented on one sheet of paper, with five categories ranging over the qualitative scale from "most characteristic" to "least characteristic." After the experimenter (T.E.H.) gave appropriate verbal instructions that differentiated the two concepts of self, the *S*s sorted the statements into the five categories, first according to their self concept and then according to their ideal self concept. The number of statements to be placed in each category were 6, 25, 38, 25, and 6, respectively, thus forcing the items into a binomial distribution.

The last step in the testing procedure was the administration of the California Test of Personality, Secondary Series [10], appropriate for students from the ninth grade to the college level. Two measures of adjustment, personal and social, are given by this test. The total score may be regarded as an indicator of over-all adjustment.

The self concept and the ideal self concept were correlated (Pearsonian r) and the resulting coefficients were transformed into Fisher's z to facilitate statistical treatment. The individual z values were then correlated with the percentile scores obtained from the California test.

The first and the fourth quartiles, each comprising 20 *S*s, were compared with respect to specific areas and specific items of the California test.

Results and Discussion

The individual correlations between self concepts and ideal self concepts ranged from —.34 to .79; the corresponding z values were from —.35 to 1.07. The mean z was .42 and the standard deviation was .26. A comparison of the z distribution with the normal curve, by the area method, yielded a χ^2 of 4.89 for 6 df; the corresponding p value was higher than .50. Consequently, the distribution of z values can

[1] For the sake of brevity, the correlation between self concept and ideal self concept will be referred to as "self-ideal congruence." Both concepts of self are operationally defined by the results of respective *Q* sorts.

be considered normal. Retransformation into Pearson r's gave a sample mean of .40. The fiducial limits at the .01 level for the mean of the universe were .33 and .46, and the lower limit of .33 at the .001 level was still significantly greater than zero. It may be concluded that the correlations between the self and the ideal self were distributed normally around a positive mean value.[2]

The correlations between self-ideal congruence (z) and the percentile scores of the California test and its subscales are given in Table 1, together with the correlation with IQ. The right-hand columns of Table 1 show a comparison between two extreme groups with regard to self-ideal congruence, i.e., *S*s in the first and fourth quartiles of the z distribution. With one exception, the correlations between self-ideal congruence and the adjustment scores are significant at least at the .05 level. The differences between the means of the two extreme groups are also significant. It is not surprising that the correlation involving "social standards" (subscore 2A) is not significant, for it is indeed doubtful whether mere knowledge of social standards reflects adjustment in general or, more specifically, social adjustment.

Table 1

Correlations between Self-Ideal Congruence, Adjustment Scores, and Intelligence

Variable	Correlation with self-ideal congruence (z)	Mean score: High self-ideal congruence	Mean score: Low self-ideal congruence	t
Total adjustment	.70*	84.50	34.25	9.29*
1. Self adjustment	.70*	80.50	36.25	6.57*
A. Self reliance	.48*	73.70	47.50	4.09*
B. Sense of personal worth	.50*	80.00	43.30	4.95*
C. Sense of freedom	.40*	69.00	36.75	4.78*
D. Feeling of belonging	.55*	75.50	40.25	5.65*
E. Freedom from withdrawal tendencies	.56*	77.00	29.40	6.00*
F. Freedom from nervous symptoms	.39*	74.50	46.65	3.42**
2. Social adjustment	.59*	80.25	33.00	7.13*
A. Social standards	.08	70.25	60.05	1.20
B. Social skills	.52*	70.90	30.00	5.59*
C. Freedom from antisocial tendencies	.22***	75.45	51.40	2.71**
D. Family relations	.30**	85.75	57.60	3.24**
E. School relations	.45*	65.90	31.50	4.53*
F. Community relations	.50*	76.75	27.20	5.33*
IQ	.05	105.80	105.95	0.05

* Significant at the .001 level.
** Significant at the .01 level.
*** Significant at the .05 level.

The F test of linearity [11, p. 278] was applied to the correlation between self-ideal congruence and total adjustment (.70), and the resulting value of .86 was found to fall far short of that required for the significance of deviations from rectilinearity at the .05 level, i.e., an F of 2.14 for 7 *df* and 69 *df*. Thus, the correlation between self-ideal congruence and total adjustment is not only marked but also rectilinear. It is concluded, therefore, that the congruence between the self concept and the ideal self concept can be used as a measure of adjustment with a considerable degree of confidence.[3] This confirmation of the basic hypothesis of the present study also lends justification to the manner in which the Rogerian school has employed correlations between self-sorts in the investigation of the psychotherapeutic process.

The correlations between IQ and self-ideal congruence (.05), IQ and total adjustment (.12), IQ and self adjustment (.11), and IQ and social adjustment (.11) were all found to be not significantly different from zero. These findings indicate that self-ideal congruence and adjustment are not a function of intelligence. This result is in substantial agreement with the findings of Kimber [3], who has reported a correlation of .04 between the Army Alpha scores and California test scores of 400 college students.

Although there was a slight tendency for

[2] This statement is also true of the distribution of Pearson r's, which is negatively skewed, but not significantly deviant from normality ($\chi^2 = 8.34$ for 5 *df*; $p > .10$). However, for purposes of averaging and subsequent correlational analysis, Fisher's z's are preferable.

[3] The regression equation between the r values of self-ideal congruence and the percentile scores for total adjustment (TA) reads: $TA = 26 + 87\ (r) \pm 17.3$. The possible values of r which satisfy this equation thus range from —.50 to 1.00.

the older *S*s to score lower on measures of adjustment and self-ideal congruence, *t* tests showed that no significant relationship was present.

A comparison of the *S*s in the first and fourth quartiles of the *z* distribution (self-ideal congruence) was made on the basis of their responses to the 180 items of the California test. The two groups showed no difference in their manner of responding to 12 of the items. For the remaining items, the low self-ideal congruence group have a higher percentage of unfavorable responses to 155, while the high self-ideal congruence group gave relatively more unfavorable responses to only 13. In the latter instance, none of the differences was statistically significant. On the other hand, of the 155 items which showed a greater percentage of unfavorable responses for the low self-ideal congruence group, 42 of the differences were significant at the .05 level and 14 at the .01 level.

A qualitative analysis of the items that discriminated between the two groups showed that *S*s in the low self-ideal congruence group were characterized by a lack of persistence in work habits and by feelings of oppression and inadequacy. They were more likely to feel rejected by both family and friends and tended to react to this by withdrawing within themselves. For this reason, their fantasy life seemed to be more active than that of the high self-ideal congruence group. The social standards and tendency to avoid antisocial activities were the same for both groups, but *S*s of the low self-ideal congruence group showed much more uneasiness in social situations. They tended to stay in the background and to avoid contacts calling for poise and diplomacy, for they were easily hurt and were aware of their proneness to quarrel. They tended to dislike school, and their negativistic attitudes made for difficulties in their relationships with teachers, fellow students, and members of the opposite sex.

It is clear from the above description that the personality features of psychotherapy patients are likely to coincide with the characteristics of *S*s showing a low congruence between self concept and ideal self concept.

Summary

The relationship between measures of adjustment and the congruence of the self concept and ideal self concept in a sample of 78 high school students was investigated, using the California Test of Personality and a modified *Q*-sort technique. The following conclusions were drawn:

1. The correlation between the self concept and the ideal self concept tends to be positive, with a mean value in an unselected population between the fiducial limits of .33 and .46. Therefore, the over-all tendency is toward a congruence of the two concepts of self.
2. Congruence between self and ideal self is a normally distributed trait.
3. The correlation between self-ideal congruence and total adjustment is positive and highly significant with regression being rectilinear. Therefore, the use of measures of self-ideal congruence in evaluating the extent of personality maladjustment appears justified.
4. The hypothesis which underlies the use of *Q* sorts in evaluating change in psychotherapy is confirmed.
5. Intelligence and age show no significant relationship with self-ideal congruence and with measures of adjustment.
6. Maladjustment in a person need not require that his self concept be negatively related to his ideal self. Where the correlation is minimal (r less than .27), signs of maladjustment may already be manifest.

Received September 8, 1953.

References

1. Bills, R. E., Vance, E. L., & McLean, O. S. An index of adjustment and values. *J. consult. Psychol.*, 1951, **15**, 257-261.
2. Draffen, Mary T. The relation between self-self ideal consistency and inconsistency and personality traits. Unpublished master's thesis, The Catholic Univer. of America, 1951.
3. Kimber, J. A. The insight of college students into the items on a personality test. *Educ. psychol. Measmt*, 1947, **7**, 411–420.
4. Mowrer, O. H. *Q*-technique—description, history and critique. In O. H. Mowrer (Ed.), *Psychotherapy: theory and research*. New York: Ronald, 1953. Pp. 316-375.

5. Raimy, V. C. Self reference in counseling interviews. *J. consult. Psychol.*, 1948, **12**, 153–163.
6. Rogers, C. R. Perceptual reorganization in client-centered therapy. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: an approach to personality*. New York: Ronald, 1951. Pp. 307-327.
7. Rogers, C. R. *Client-centered therapy*. Boston: Houghton Mifflin, 1951.
8. Rogers, C. R. Studies in client-centered psychotherapy III: the case of Mrs. Oak — a research analysis. *Psychol. Serv. Center J.*, 1951, **3**, (1-2), 47-165.
9. Seeman, J., & Raskin, N. J. Research perspectives in client-centered therapy. In O. H. Mowrer (Ed.), *Psychotherapy: theory and research*. New York: Ronald, 1953. Pp. 205-234.
10. Tiegs, E. W., Clark, W. W., & Thorpe, L. P. The California test of personality. *J. educ. Res.*, 1941, 35, 102–108.
11. Walker, H. M., & Lev, J. *Statistical inference*. New York: Holt, 1953.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Validity of a Critical-Item Scale for Schizophrenia on the MMPI[1]

Harold Rubin
VA Hospital, Lebanon, Pa.

In an attempt to eliminate the problem of "false positives" which confronts many clinicians using the MMPI, Benarick [1] has developed a critical-item scale to aid in differentiating psychotics from nonpsychotics. The scale, reported earlier in this journal [2], consists of 11 items taken from the *Sc* scale of the MMPI and scored in the *Sc*-plus direction. Benarick also presents a second scale consisting of 10 items answered more frequently in the *Sc* direction by nonpsychotics. The present study attempts to test the validity of these critical-item scales for schizophrenia on a hospital population. It differs somewhat from Benarick's original study in that the nonpsychotic group consists of hospitalized subjects.

Procedure

MMPI profiles selected were those of 64 male patients of a veterans hospital who were administered the MMPI as part of an admission test battery, and whose raw *Sc* score was greater than 20. Thirty-two of the patients were diagnosed as schizophrenics with a mean raw *Sc* score of 29.6, and thirty-two were diagnosed as psychoneurotics with a mean raw *Sc* score of 28.3. All diagnoses were made by the psychiatric staff of the hospital following a period of observation.

Results and Discussion

As may be observed in Table 1, the mean scores on the eleven-item scale were 3.8 for the psychotics and 2.9 for the nonpsychotics. The highest number of critical items obtained by any single subject was eight, this being true in both groups.

Following Benarick's [2] suggestion, a cutting score of 4 was used to identify psychotics from nonpsychotics. With this criterion, it was found that 59% of the psychotics were correctly identified, while 31% of the nonpsychotics yielded "false positives." For the total group of 64 patients, 62% of the identifications made were correct. In an earlier study, Rubin [4] has shown that psychotics tend to have their highest scores on the *Sc* scale. On this basis, if one were to assume that each profile represented that of a psychotic, 50% of this group of 64 would be correctly identified. Benarick's scale offers a 12% improvement in such correct identifications.

A chi square computed between psychotics and nonpsychotics, using the cutting score of 4, yielded a value of 5.11, which is significant at less than the 5% level. Although not as significant as the findings reported by Benarick [2], some validity, it seems, may be attributed to this critical-item scale. However, when chi-square values were computed separately for each of the eleven items, not a single item was

Table 1

Distributions of Scores of Psychotics and Nonpsychotics on Two Critical-Item Scales

Score	Eleven-Item Scale		Ten-Item Scale	
	Psychotic	Nonpsychotic	Psychotic	Nonpsychotic
10				
9			1	2
8	1	1	1	1
7	2	0	5	7
6	2	1	7	4
5	7	5	6	8
4	7	3	6	5
3	5	8	2	2
2	3	6	3	2
1	5	6	1	1
0	0	2	0	0
Total	32	32	32	32
Mean	3.8	2.9	5.0	5.3

[1]From the Veterans Administration Hospital, Lebanon, Pennsylvania.

found to differentiate significantly between psychotics and nonpsychotics. The distribution of scores among the eleven items was markedly similar for both groups. One exception was that on three of the items (76, 157, and 331) the nonpsychotics obtained a slightly greater, although not significant, frequency. This is in contrast to the findings of Benarick, who reported significant differences between psychotics and nonpsychotics on the individual items as well as on the total scale.

On the ten-item scale, as observed from Table 1, it was found that the distributions for both psychotics and nonpsychotics were so similar that further statistics were unnecessary. It should be stated, however, that among both psychotics and nonpsychotics there was a tendency for those who obtained a high score on the eleven-item scale to obtain a high score on the ten-item scale. This would indicate that the scales are not as complementary as one might expect from Benarick's results.

The findings of the present study would appear to indicate that even though it was possible to combine a number of items into a scale having some diagnostic value when used on a group level, the items taken singly failed to be statistically meaningful in separating psychotics from nonpsychotics. Since only four of the critical items need be scored to make a differential diagnosis or to detect "false positives," which, after all, is the purpose of the scale, it would seem that the value of the scale in differentiating between psychotics and nonpsychotics is limited, particularly when the subjects are hospitalized patients. The present study also indicates the difficulty of constructing a meaningful scale based on psychiatric diagnoses which are themselves of questionable reliability [3].

Received September 30, 1953.

References

1. Benarick, S. J. An investigation of the schizophrenic scale of the Minnesota Multiphasic Personality Inventory. Unpublished master's thesis. The Pennsylvania State College, 1950.
2. Benarick, S. J., Guthrie, G. M., & Snyder, W. U. An interpretive aid for the *Sc* scale of the MMPI. *J. consult. Psychol.*, 1951, **15**, 142-144.
3. Rotter, J. B. Clinical methods: psychodiagnostics. *Annu. Rev. Psychol.*, 1953, **4**, 295-316.
4. Rubin, H. The MMPI as a diagnostic aid in a veterans hospital. *J. consult. Psychol.*, 1948, **12**, 251-254.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Wechsler-Bellevue Performance of Delinquent and Nondelinquent Girls

Julia R. Vane
Hempstead Public Schools
and Virginia W. Eisen
Catholic Charities Guidance Institute

The purpose of this study was to determine whether a group of adolescent female delinquents would show a characteristic test pattern on the Wechsler-Bellevue Scale. Other studies concerned with delinquent patterning have dealt with male offenders, with one exception, a recent study by Bernstein and Corsini [4]. These authors investigated the effect of order of presentation of the subtests of the Wechsler-Bellevue Scale on the scores of female delinquents. Their hypothesis that the psychopathic pattern described by Wechsler is a function of the conditions under which delinquents are tested was not substantiated. They found that the rank order of the subtests was virtually the same for the two groups despite the difference in order of presentation.

The subjects in the present study included an experimental group of 100 delinquent females between the ages of 16 and 25, with IQ's between 70 and 110 on the Wechsler-Bellevue Scale, and a control group of 38 nondelinquent females in the same age and IQ range. All *S*s in both groups were white and native born. All were interviewed and given the Wechsler-Bellevue Scale by one of the authors.

The experimental group consisted of delinquents committed by the New York State courts to the House of Good Shepherd, Peekskill, New York, a reformatory for girls and women adjudged guilty of offenses such as prostitution, forgery, and shoplifting. Psychiatric and psychological interviews and case history material were used to eliminate delinquents suffering from psychoses, brain damage, alcoholism, or psychoneuroses. The great majority of the delinquents (93%) fell within the age range in which the greatest incidence of crime is reported, i.e., 16 to 21 years [8], and in this respect should be representative of delinquents in general.

The control group consisted of girls from two New York City public high schools. To avoid inclusion of possible delinquents, no girl was selected who was considered a behavior, personality, or social adjustment problem. Truancy, even of the shortest duration, was also a basis for exclusion, since it had been found that 99% of the delinquent group had records of truancy. School guidance cards, attendance records, and interviews with the school guidance counselor, as well as interviews with the girls themselves, were used to evaluate the variables.

The nondelinquents were equated as closely as possible with the delinquents for level of intelligence and chronological age. Years of schooling could not be matched since in order to obtain nondelinquents similar in age to the delinquents, it was necessary to choose the nondelinquents from the junior and senior years of high school. As the result all the nondelinquents had completed at least three years of high school, whereas 91% of the delinquents had left school during the first and second years of high school, if not before. Practically all of the nondelinquents were enrolled in vocational or nonacademic courses.

With respect to socioeconomic background, the delinquents and nondelinquents closely resembled one another. The majority of the girls from both groups resided in the same areas of New York City, and their fathers' occupations were similar. On the Goodenough-Anderson Occupational Rating Scale [6] 82% of the fathers' occupations in both the delinquent and the nondelinquent group fell in the lower three categories, namely, V—semiskilled, minor clerical, minor business, VI—slightly skilled, and VII—day laborers.

The majority of delinquents (53%) were the

children of parents both born in the United States. Of the remainder, not more than 10% had parents who were born in any one particular country. Half of the nondelinquents (50%) also had both parents born in the United States. Of the remainder, 26% were the offspring of parents one or both of whom were born in Italy. No other country was the birthplace of more than 8% of the parents of the nondelinquents.

With respect to intelligence, the mean total IQ of the control group showed a slight superiority, but the difference was not statistically significant ($t = 1.14$).

The mean chronological age of the delinquents (18.5) was 1.2 years greater than the mean age of the nondelinquents (17.3). Despite the statistical significance of this difference ($t = 6.00$), it should not influence the results since Wechsler grouped 17-, 18-, and 19-year-olds into one unit in standardizing the test.

Procedure

Mean verbal, performance, and full IQ's, and mean weighted subtest scores were computed for each group, and scatter was determined in terms of these. Because of the narrow age range and the fact that the groups were equated for level of intelligence, it appeared unnecessary to convert weighted scores into z scores as suggested by Barnett [3] or into subtest IQ scores as suggested by Alimena [1]. Rather extensive experimentation was done with both methods and neither was found useful or of greater value than the use of weighted subtest scores.

Results and Discussion

Table 1 shows the results for the two groups. The largest subtest difference was the superiority of the nondelinquents over the delinquents in Digit Symbol ($t = 3.88$). Wechsler [9] has noted that unstable individuals do poorly on this test and attributes this to their difficulty in concentrating and applying themselves for any length of time to a task requiring persistent effort. It is the authors' experience that delinquent girls do show inability for sustained effort, especially in simple, relatively monotonous tasks such as the Digit Symbol test. They often verbalize resistance to such tasks. The delinquents in Glueck and Glueck's investigation also showed a significantly poorer score on Digit Symbol than the nondelinquents. Glueck and Glueck feel that the Digit Symbol test "involves concentration and persistency of effort, and is therefore a test in which impulsive, emotionally labile persons would not perform as well as those who are stable" [5, p. 207].

Table 1

Wechsler-Bellevue Scores of Two Groups

Item	Delinquents ($N = 100$)		Nondelinquents ($N = 38$)		
	Mean	*SD*	Mean	*SD*	*t*
CA	18.5	1.46	17.3	.89	6.00
Weighted score	9.12	1.10	9.37	1.02	1.25
Full IQ	95.2	8.40	96.8	6.92	1.14
Verbal IQ	93.6	9.72	93.1	7.68	.31
Performance IQ	98.5	10.08	100.9	8.92	1.35
Information	8.43	2.22	9.39	1.97	2.46
Comprehension	9.31	2.17	9.47	2.19	.31
Digits	8.24	2.42	7.53	1.67	2.03
Arithmetic	6.48	2.38	6.61	2.43	.28
Similarities	8.15	2.22	8.08	2.26	.16
P. Arrangement	10.16	2.42	10.18	2.70	.00
P. Completion	9.64	2.43	9.26	2.17	.89
Block Design	9.70	2.83	9.97	2.13	.16
Object Assembly	10.86	2.43	11.24	1.87	1.00
Digit Symbol	10.42	2.55	11.74	1.43	3.88

The second largest difference between the groups was on the Information subtest ($t = 2.46$). On this, the nondelinquents were also superior. Although this may be related to the fact that our delinquent girls had fewer years of schooling than the nondelinquents, it seems to be of diagnostic significance. It has been noted repeatedly by educators, psychologists, and court workers that delinquents rarely work to capacity in school and consequently experience failure and frustration. They often seem to try to relieve their feelings of frustration by truancy. The comparatively low Information score of the delinquents seems to be diagnostic in that it is a reflection of that aspect of the delinquent personality which operates to prevent him from taking advantage of his educational opportunities.

On the Digit Span subtest the delinquents were superior to the nondelinquents ($t = 2.03$). The difference between the two groups is actually due to the fact that the nondelinquents did poorly on this subtest, and not that the delinquents did well. This is indicated by the fact that the mean score of the nondelinquents on this test (7.53) is considerably below their mean score on the ten subtests (9.37) and is also considerably below Wechs-

ler's standardization subtest mean (9.3) for this age group. It is difficult to explain this result. It is possible that the testing situation was more anxiety provoking to the nondelinquents than to the delinquents and that this anxiety served to lower the Digit Span scores of the former, but actually it was not noted that the nondelinquents appeared any more anxious than the delinquents. It is possible that this should not be considered a valid indicator until further verification is obtained.

On the Comprehension test there was very little quantitative difference between the delinquents and the nondelinquents, but there were qualitative differences. The delinquents frequently responded to the questions by statements such as "Do you want me to tell you what I *should* do or what I *would* do?" They occasionally spoiled a good answer to the question "What is the thing to do if you find an envelope in the street . . . " by adding a remark such as "I would probably open it to see if there was any money in it." Asked "Why are laws necessary?" several countered with "You tell me!" or to the question why a marriage license is required they replied "I wish I knew." Several indicated that they would notify an usher or manager if they were the first person to discover a fire while sitting in the movies, and then added "I would really probably run." In these replies we see that the delinquents are aware of the socially accepted standards, although they may fail to apply them to their own behavior.

In both the delinquent and nondelinquent groups, the Performance IQ was considerably higher than the Verbal IQ, the difference being statistically significant (delinquents, $t = 3.5$, and nondelinquents, $t = 4.1$). As previous investigators [4, 5, 7, 9] had considered superiority of Performance scores over Verbal scores a delinquent or psychopathic sign, the fact that our nondelinquent group showed this superiority was somewhat unexpected.

Several hypotheses were considered as possible explanations for this result. One was that because the control group was more heavily weighted with *Ss* whose parents were born in Italy (26%) than is the normal population, this factor might have influenced the results. It has been found that on certain other tests "Italian children tend to rate significantly higher on the performance scale (Arthur Point) than on the Binet" [2, p. 6]. To check this hypothesis the scores of the nondelinquents of Italian background were compared with the scores of the nondelinquent group as a whole. It was found that the mean Verbal IQ of the girls with one or both parents born in Italy was 91.9, and the mean Performance IQ was 99.7. The difference between the two scores was 7.8, which is exactly the difference found between the mean Verbal and Performance IQ's of the total nondelinquent group. In addition, the percentage of girls with Italian-born parents whose Performance score exceeded their Verbal score by five or more points was compared with the percentage of girls of non-Italian born parents whose Performance score was superior by five or more points. No significant difference was found. It was therefore felt that the fact that the control group contained a large percentage of girls whose parents were born in Italy did not account for the result.

A second hypothesis was that the mean IQ's of both groups in the present study, four to five points lower than Wechsler's standardization mean IQ of 100, might have had some bearing on the difference between Verbal and Performance scores, i.e., that lower IQ groups may normally tend to do better on Performance than on the Verbal scale. To check this explanation, both the delinquent and nondelinquent groups were divided into two subgroups on the basis of high and low IQ. The superiority of Performance score over Verbal score remained statistically significant in each case, although the difference was somewhat smaller in the higher IQ subgroups.

A further possibility is that since our two groups belonged to a lower socioeconomic level than the population in general, this may have influenced the results. As we had no nondelinquent group of higher economic level, and as figures have not been reported from any comparable group, this hypothesis cannot be tested at the present time.

The most plausible conclusion seems to be that although Performance superiority frequently has been associated with delinquency, actually this may be a characteristic pattern of normals at this age level. The fact that Wechsler's standardizing data [9, p. 222] for age levels 16 to 19 years also show a significant superiority (approximately $t = 3.0$) of Performance scores over Verbal scores would appear to substantiate this conclusion.

It should be noted in addition that two other signs associated with delinquency are present in our nondelinquent group. The sign "Picture Arrangement plus Object Assembly greater than Block Design plus Picture Completion" is found in 63% of the delinquent group

and 66% of the nondelinquent group, and Arithmetic is the lowest ranking of the subtests in each group. The fact that the three above-mentioned signs have been associated with delinquents [4, 5, 7, 9] may have arisen because the delinquents were not compared with a sufficiently similar group of nondelinquents. It has been demonstrated that these signs are found in delinquent test profiles, but our results show that they are not unique for delinquents and occur also in certain normal groups.

In general it would seem that the successful use of test signs in differential diagnosis varies with the circumstances. For instance, it may be possible to differentiate delinquents from other clinical groups, such as neurotics, psychotics, and organics, on the basis of their Wechsler-Bellevue profiles, but these signs would not be adequate to differentiate between the "normal" nondelinquent girl and the true delinquent. It would appear that the act of delinquency itself would be a sufficient means of differentiation. Our experience has shown, however, that this is not always a good criterion, for the reformatory has girls who cannot be considered truly delinquent. Some have been committed by punitive families for a single sexual indiscretion. Others who ran away from intolerable home conditions were committed "to keep them out of trouble"; girls of this type were not included in our delinquent group. In contrast to these, the records of demonstrated psychopaths often showed previous suspended sentences on more serious charges. Actually, our records indicate that truancy is a highly crucial factor in the picture of delinquency and in the history of the individual delinquent. Serious consideration of this factor and more adequate and psychologically oriented methods of handling the early truant would thus appear important.

While our results do show some quantitative and qualitative test differences between the delinquent and the nondelinquent, they still seem to fall short of the desired accuracy. It would appear that the Wechsler-Bellevue Scale is not an adequate means of differentiating delinquents from nondelinquents when age, intelligence, and socioeconomic background are similar.

Summary

A group of delinquent and a control group of nondelinquent girls, equated for age and intelligence and similar in socioeconomic background, were given the Wechsler-Bellevue Intelligence Scale to determine if the delinquents would give a characteristic pattern on this test.

There was a statistically significant difference between these two groups on only two of the subtests, viz., Information and Digit Symbol, the delinquents being inferior to the nondelinquents on both. On the whole, the delinquents and the nondelinquents were more alike than different, showing similar deviations from Wechsler's norms for their age group. Three of the signs offered by Wechsler and Levi [7, 9] as characteristic of male psychopaths were found in both the delinquents and the nondelinquents in the present study.

On the basis of present findings it is concluded that female delinquents do show certain deviations from the standardization norms in respect to subtest variation on the Wechsler-Bellevue Scale. These deviations cannot, however, be considered unique to the delinquent since they also appear in nondelinquents of similar age, socioeconomic, and educational level. One might hypothesize that age, socioeconomic, educational, and cultural factors play a larger part in determining qualities of test performance on the Wechsler-Bellevue Scale than do deviant personality factors, but this hypothesis has yet to be adequately demonstrated.

Received October 16, 1953.

References

1. Alimena, B. Norms for scatter analysis on the Wechsler intelligence scales. *J. clin Psychol.*, 1951, 7, 289–290.
2. Arthur, Grace. *A point scale of performance tests.* Vol. 1. *Clinical manual.* (2nd Ed.) New York: The Commonwealth Fund, 1943.
3. Barnett, I. The use of z-scores in equating the Wechsler-Bellevue subtests. *J. clin. Psychol.*, 1950, 6, 184–188.
4. Bernstein, Rachel, & Corsini, R. J. Wechsler-Bellevue patterns of female delinquents. *J. clin. Psychol.*, 1953, 9, 176–179.
5. Glueck, S., & Glueck, Eleanor. *Unraveling juvenile delinquency.* New York: The Commonwealth Fund, 1950.

6. Goodenough, Florence, & Anderson, J. E. *Experimental child study.* New York: Century, 1931.
7. Levi, J. A psychometric pattern of the adolescent psychopathic personality. Unpublished doctor's dissertation, New York Univer., 1943.
8. Tappan, P. W. *Juvenile delinquency.* New York: McGraw-Hill, 1949.
9. Wechsler, D. *The measurement of adult intelligence.* (3rd Ed.) Baltimore: Williams & Wilkins, 1944.

Journal of Consulting Psychology
Vol. 18, No. 3, 1954

Note on Sanders and Pacht's Investigation of Size Constancy of Clinical Groups

Edward Lovinger [1]
Michigan State College

In an investigation of perceptual size constancy of clinical groups, Sanders and Pacht [2] employed a rather unusual size-constancy index. The purpose of this note is to point out the differences between their index and those more commonly utilized and to demonstrate how its use is likely to be misleading.

They reported mean constancy indices of 105 for a group of normals, 99 for a group of neurotics, and 88 for a group of psychotics. The technique they employed for measuring constancy was one in which the comparison stimuli were far and the standard near.[2] The comparison stimuli consisted of 13 black squares graduated in size and randomly arranged under a consecutive series of numbers. These squares were always present at a distance of five meters from the subject. The standard squares were of five different sizes and were presented individually at two distances from the subject, first at four meters and then at two meters. The subjects were instructed to choose the comparison square that was equal in size to the standard square.

The following equation was utilized in obtaining the size-constancy index:

$$\text{Size-constancy index} = \frac{\text{Size}_{near} - \text{Size}_{far}}{\text{Size}_{near}}.$$

Size_{near} and Size_{far} refer to the judged size of the standard at two and four meters respectively. In order to eliminate negative scores and decimals, a constant of one was added to the obtained ratio, and this result was multiplied by 100. Thus, perfect constancy would be represented by an index of 100. An index greater than 100 would indicate that the comparison square selected for the far standard (four meters) was smaller than the comparison square selected for the near standard (two meters), i.e., that the judged size of the standard was smaller when it was far than when it was near. An index of less than 100 would indicate that the judged size of the far standard was larger than that of the near.

Although this method of computation results in an index which clearly indicates the comparative judged sizes of the standard at the two distances, this index is likely to lead to misinterpretations if it is accepted as equivalent to a Brunswik or Thouless ratio. Sanders and Pacht at one point in their discussion seem to have done this. The authors stated: "The present results suggest findings which are similar to those of Thouless . . . , namely, that psychotics show less size constancy than normals" [2, p. 443]. This statement suggests not only a seeming misinterpretation of their own data, but also a misunderstanding of Thouless' experiment. Thouless [3] investigated the differences in constancy between cyclothymes and schizothymes. He found no significant differences between the two groups for three different constancy measures, although all the insignificant differences were in the direction of less constancy (underconstancy) for the schizothymes. Since there was no indication that any of these subjects were psychotic, including the schizothymes, it would seem rather unwarranted to speak of Thouless' results as indicating less size constancy for *psychotics* than for normals.

More important, however, is the implication that the authors' constancy index is equiv-

[1] The author wishes to thank Drs. M. Ray Denny and Paul D. Greenberg for their assistance and suggestions. However, the author assumes sole responsibility for the content of the note.

[2] Note that the terms *standard* and *comparison* have been substituted for Sanders and Pacht's *stimulus squares* and *fixed standard series*, respectively, in order to conform to the terminology generally found in the literature.

alent to a Thouless or Brunswik ratio. It seems apparent from the statement quoted above that Sanders and Pacht are treating the mean constancy index of 88 for the psychotics as equivalent to Thouless' ratio, and hence they consider their findings similar to those of Thouless. It will be shown presently, however, that their index of 88 actually indicates overconstancy rather than underconstancy. Hence, their results for the psychotics are opposite rather than similar to Thouless' findings for schizothymes. The fact that their results differ from those of Thouless is not so important as is the fact that their size constancy index signifies deviations from perfect constancy (100 for Sanders and Pacht and 1.0 for Brunswik or Thouless ratios) which are opposite in direction from the usual size constancy index. For example, a Brunswik or Thouless ratio of less than 1.0 indicates underconstancy, i.e., a judgment which falls somewhere between retinal proportions and real object proportions. An index greater than 1.0 indicates overconstancy, i.e., a judgment which goes beyond real object proportions in a direction opposite from retinal proportions.

Sanders and Pacht point out that the psychotic group's size judgments go beyond real object proportions in a direction away from retinal proportions. Hence, this would indicate that their mean constancy index of 88 indicates not underconstancy but overconstancy. Furthermore, this becomes even more apparent when we consider the relative sizes of the matches for the standard at the two distances. Underconstancy would be indicated when a subject's match for the standard at a distance of four meters is smaller than his match at two meters since the size of the retinal image of an object decreases with distance. However, for these matches Sanders and Pacht's formula would have given an index of more than 100. When the subject's judgment for the far standard is larger than that for the near standard (overconstancy), the resultant Sanders and Pacht index would be less than 100. This inverse relation to the Brunswik or Thouless ratios becomes most apparent when one examines a hypothetical situation wherein our subject's judgments conform strictly to retinal proportions. With Sanders and Pacht's experimental arrangement, when the standard is 50 mm. in size, our subject's judgments would be 125 mm. for the near standard and 62.5 mm. for the far standard. Such a situation would result in Brunswik and Thouless ratios of zero, whereas Sanders and Pacht's index would be 150.

If one does not take into account the discrepancy between Sanders and Pacht's constancy index and Brunswik or Thouless ratios, much of their discussion seems confusing and appears to contradict their results. For example, they point out that their data indicate a breakdown in constancy for schizophrenics opposite in direction from that predicted by Bruner [1]; i.e., the schizophrenics' judgments, rather than conforming more with retinal proportions, go beyond real object proportions in a direction opposite from retinal proportions. This statement seems to contradict their report of a mean constancy index of 88 for the psychotic group when this index is accepted as equivalent to a Brunswik or Thouless ratio. However, when the fact that their index indicates deviations from perfect constancy in a direction opposite from Brunswik or Thouless ratios is taken into account, then the seeming contradiction disappears.

It is recognized that the usual methods of computation would not have given a single size-constancy index for the standard at two positions, but rather, two indices, one for each position of the standard. However, if the authors desired a single index, they could have obtained indices comparable to the usual ones simply by altering their equation to read as follows:

$$\text{Size-constancy index} = \frac{\text{Size}_{far} - \text{Size}_{near}}{\text{Size}_{far}}.$$

Just as for their equation, a constant of one should be added.

One wonders why, in the interest of consistency with previous research in this area, the authors chose an index not equivalent to those commonly used.

Received September 18, 1953.

References

1. Bruner, J. S. Personality dynamics and perceiving. In R. Blake, & G. Ramsey (Eds.), *Perception — an approach to personality.* New York: Ronald, 1951. Pp. 121-145.
2. Sanders, R., & Pacht, A. R. Perceptual size constancy of known clinical groups. *J. consult. Psychol.*, 1952, **16**, 440-444.
3. Thouless, R. H. Individual differences in phenomenal regression. *Brit. J. Psychol.*, 1932, **22**, 216-241.

NEW BOOKS AND TESTS

Books

Bach, George R. *Intensive group psychotherapy.* New York: Ronald, 1954. Pp. xi + 446. $6.00.

Intensive Group Psychotherapy attempts a marriage of group therapy and Kurt Lewin's group dynamics. Based on two continuous intensive therapy groups and liberally interspersed with protocols and brief examples from the therapy sessions, the volume is rich in its formulations of the many factors involved in the group therapy process. Its weakness is that, if anything, it is overorganized, possibly giving the student in group therapy the feeling that therapy can be done by rule of thumb methods. There are, for example, 4 personality criteria for exclusion, no more, no less, 16 communication themes, 7 therapeutic functions in role playing, 7 phases of "theragnostic work," 18 do's and 10 don'ts, 7 developmental phases, 7 communicational levels, 7 instigators of defensive resistance, and 8 factors of group dynamics. In spite of the mechanistic organization, this volume is one of the most intriguing works on group therapy to appear in some time. It is a brave attempt to bring order out of chaos, even if it is rigidly structural. If one accepts the author's axioms, the conclusions follow logically.—*M.K.*

Christie, Richard, & Jahoda, Marie. (Eds.) *Studies in the scope and method of "The authoritarian personality."* Glencoe, Ill.: Free Press, 1954. Pp. 279. $4.50.

In this unusual volume, seven authors make a retrospective evaluation of *The Authoritarian Personality* (1950) by examining some of its concepts, its methods, and the subsequent studies it inspired. Shils and Laswell contribute essays that deal with the study's structure of concepts and theory. In perhaps the most telling chapter, Hyman and Sheatsley re-examine problems of method with respect to sampling, measuring instruments, and methods of analysis. Christie reviews the newer research studies. One salient point that emerges both from the critique and from subsequent research is the need to control education and social class in studies of authoritarianism. Some of the attributes and correlates of the trait now seem to be identified as social class differences. In a concluding section, Frenkel-Brunswik reviews the study from the viewpoint of an original contributor, and reports new data from an investigation of children's prejudices. This critical volume gives a global impression that *The Authoritarian Personality* possessed many faults of method and hence of conclusions, but that it opened to investigation a significant area hitherto dominated by opinion and controversy.—*L. F. S.*

Eissler, Ruth S., *et al.* (Eds.) *The psychoanalytic study of the child.* Vol. VIII. New York: International Universities Press, 1953. Pp. 412. $7.50.

This eighth volume of a now well-established series is up to the excellent standard of previous volumes. The nonanalytic child psychologist should not be frightened by its theoretical orientation, as the case material is extensive and worth reading in its own right. The result is a mature and stimulating symposium.—*W. A. H.*

Eysenck, H. J. *Uses and abuses of psychology.* London & Baltimore: Penguin Books, 1953. Pp. 318 (paper). 65¢.

In this smoothly and interestingly written nontechnical book, Eysenck discusses intelligence testing, vocational psychology, abnormal behavior, and social attitudes. As might be predicted from his other writings, the author emphasizes the findings from research and takes a dim view of intuitive appraisals, psychoanalytic theory, and psychotherapy in general.—*L. F. S.*

Hartley, Eugene L. (Ed.) *Doubleday papers in psychology.* Garden City, N. Y.: Doubleday, 1954.

Leuba, Clarence. *The natural man.* Doubleday Papers in Psychology, No. 1, 1954. Pp. x + 70 (paper). 95¢.

English, Horace B. *The historical roots of learning theory.* Doubleday Papers in Psychology, No. 2, 1954. Pp. viii + 21 (paper). 65¢.

Frank, Lawrence K. *Feelings and emotions.* Doubleday Papers in Psychology, No. 3, 1954. Pp. viii + 38 (paper). 85¢.

These three pamphlets are the first of an announced series of fourteen papers on psychological topics. As judged from the initial sample, the pamphlets are clear, elementary, interesting, and at-

Note.—The reviews were prepared by the Editor and the Associate Editors, who may be identified by their initials.

tractive. An assembly of appropriate topics might provide text or reference material for a course in introductory psychology or for extension courses. The pamphlets would not be suitable for advanced or graduate instruction. Leuba's *The Natural Man,* an excellently integrated presentation of studies of chimpanzees and of nonliterate cultures, provides a novel introduction to the problems of psychology. English, in the second pamphlet, reviews theories of learning from Descartes and Hobbes to E. L. Thorndike. Frank's essay, *Feelings and Emotions,* is somewhat more diffuse and "popular" than the other two papers. It cites no references, ignores many issues that most psychologists would regard as relevant to its topic, and digresses into the areas of child development and personal adjustment.—*L. F. S.*

Ingham, Harrington V., & Love, Leonore R. *The process of psychotherapy.* New York: McGraw-Hill, 1954. Pp. ix + 270. $5.00.

This is a description of the way in which a psychotherapist works, with no attempt to introduce dynamics of personality except when necessary to understand treatment procedure. The noncontroversial conditions preliminary to, and occurring during, therapy are presented somewhat succinctly and simply. Two cases are included. One is a verbatim reproduction of treatment covering a single hour; the second describes events in 36 psychotherapeutic interviews. The absence of even a single reference to other books or periodicals is striking. There is no indication as to whom the book is addressed, but the authors might have indicated more emphatically the dependence of the technique upon an understanding of personality theory.—*F. McK.*

Libo, Lester M. *Measuring group cohesiveness.* Ann Arbor, Mich.: Univer. of Michigan Press, 1953. Pp. ix + 111 (paper). $2.00.

This research bulletin contains a discussion of the concept of group cohesiveness, a report of research on its measurement, and a new projective test for appraising attraction to group. The test shows excellent scoring reliability and significant relationships to behavioral measures of group cohesiveness but is presented for experimental use, not as a finished instrument for prediction.—*L. F. S.*

Klopfer, Bruno, Ainsworth, Mary D., Klopfer, Walter G., & Holt, Robert R. *Developments in the Rorschach technique.* Vol. I. *Technique and theory.* Yonkers, N. Y.: World Book Co., 1954. Pp. x + 726. $6.50.

Replacing the widely used *The Rorschach Technique,* by Klopfer and Kelley, this fat volume is at once a manual of administration, scoring, and interpretation, and a theoretical monograph that attempts to bring the Rorschach test within the purview of contemporary theories of personality. Probably the great majority of clinical psychologists will find something of interest and value here, a situation which attests at once to the book's comprehensiveness and to the depths to which the Rorschach has penetrated the clinical culture. One anticipates that the same kind of comment will be in order when a succeeding volume, *Fields of Application,* is published. In its role of manual, *Technique and Theory* represents a considerable advance over the older Klopfer and Kelley. There is much more detailed attention to scoring problems and to the clarification of scoring principles. The matter of interpretation is discussed with thoroughness, and with a greater degree of awareness than before that there are important validation problems still unsolved. The couching of interpretative assertions in the form of hypotheses may represent a decided advance in thinking about the place of the Rorschach in the clinician's tool kit, at once warning against dogmatism and providing an impetus for serious research on the instrument's actual utility. This part of the book is informative, teachable, and most welcome. The three chapters comprising the "theoretical monograph" are stimulating, if a bit uneven. Holt's chapter on the implications of contemporary personality theories for a Rorschach rationale is exceptionally able and constitutes the most serious and sophisticated effort to date to relate the test to broad and potentially testable issues in theoretical psychology. One senses a degree of defensiveness in Dr. Ainsworth's discussion of validation problems, but she has done a fair, intelligent, and thorough evaluation of the voluminous literature, and her chapter will be invaluable as a source of information. Bruno Klopfer's concern with Rorschach hypotheses and ego psychology is essentially speculative and given to too certain interpretative assertions. ("The earliest communicated response to shading shows a mixture of *cF* and *Fc.* This represents a considerably differentiated need for affection and belongingness" p. 590.) But read with an eye both to clinically useful insights and to the formation of hypotheses about Rorschach and about personality functioning, this vigorous chapter is challenging and evocative. There can be no gainsaying the importance of the book at this particular stage of clinical psychology's development. The judgment of history will depend on the quality and outcome of the research that it should stimulate.—*E.J.S.*

Paterson, Donald G., Gerken, C. d'A., & Hahn, Milton E. *Revised Minnesota Occupational Rating Scales.* Minneapolis: Univer. of Minnesota Press, 1953. Pp. ix + 85 (paper). $2.00.

This 85-page pamphlet is a second revision of the Minnesota Occupational Rating Scales, first offered in 1937. The scales consist of ratings, based on the pooled judgment of vocational psychologists, for seven abilities or aptitudes as these are required for

each of 432 occupations. The abilities, each rated at four levels, are academic ability, mechanical ability, social intelligence, clerical ability, musical talent, artistic ability, and physical agility (new in this revision). The variation in the ratings from scale to scale yields an occupational pattern, e.g., ratings for eminent lawyer are A D A B D D D. These ratings are presented in a table in which jobs are arranged alphabetically, each preceded by an identifying number, and followed by the D.O.T. code number for jobs at every level of each ability, and a pattern classification that makes easy a consideration of occupational alternatives. Uses of the scales and the Individual Counseling Record are explained. By this system, both the counselee and the job are rated in the same terms.—*A. R.*

Tyler, Leona E. *The work of the counselor.* New York: Appleton-Century-Crofts, 1953. Pp. xi + 323. $3.00.

This is primarily a text for a first course in counseling. The author conceives counseling as "more a matter of what one perceives or comprehends, than what one says or does," and emphasizes this position most effectively in the context of discussing such practical matters as the use of records, tests, or occupational information. Because she admits that "counseling is still to a large extent an art," she writes largely in terms of wisdom based on "shared experiences," including her own. However, each chapter includes a research summary evaluating the relevant research evidence, and the final chapter is devoted to a summary of studies on the evaluation of the outcomes of counseling. Included is an up-to-date bibliography of 175 titles. Although designed as an introductory text, this little volume might well be read with profit by experienced counselors. It deals with basic issues and treats them in a manner that this reviewer found both challenging and sensible. The style of writing is superior.—*E. L. K.*

Witkin, H. A., Lewis, H. B., Hertzman, M., Machover, K., Meissner, P. Bretnall, & Wapner, S. *Personality through perception.* New York: Harper, 1954. Pp. xxvi + 571. $7.50.

Witkin and his collaborators here report one of psychology's best achievements in the integration of theory, the laboratory, and the clinic. The major variable in their elaborate ten-year study is space orientation—the tendency of the subject to ". . . locate the upright according to the axes of the visual field or with reference to sensations from his own body. . . ." Tests of space orientation are studied in relation to one another, to other perceptual and motor tests, and to clinical appraisals, which include an interview, the Rorschach, the Figure-Drawing Test, and the TAT. Subjects are college students of both sexes, children from 8 to 17, and mental hospital patients. In the main, the data seem to confirm the hypotheses that "field dependence" in the perception of the vertical is a consistent variable, and that it has meaningful relationships to other features of the organization of personality. The studies of sex differences, and of the curvilinear relationships of certain perceptual variables with age from 8 years to adulthood, are especially penetrating. The methods for treating the data of the interviews and projective tests, achieving quantification without undue loss of clinical insights, may serve as models for future studies. Amid all these real merits, the work has one serious shortcoming. The interviews and projective devices were analyzed a posteriori by examining the performances of groups already identified as different in their space perceptions. The research with the personality tests is therefore exploratory and suffers from the classic error of discovering hypotheses and testing hypotheses with the same group of subjects. Cross validation is not entirely absent, but is weak and insufficiently stressed; procedures had to be modified for the children and for the patients in such a way as to leave indeterminate the extent to which the findings are due to the changed procedures, to the character of the samples, or to statistical shrinkage. The one genuine cross-validation group (p. 461) is small and incompletely reported. These faults can, and almost surely will, be remedied by further research studies. On the whole, the coordinated research program has great value, and the report should command the attention of all psychologists from the purest experimentalist to the most practice-centered clinician.—*L. F. S.*

Tests

Flanagan, John C. *Flanagan Aptitude Classification Tests.* Battery of 14 tests: Inspection, Coding, Memory, Precision, Assembly, Scales, Coordination, Judgment and Comprehension, Arithmetic, Patterns, Components, Tables, Mechanics, Expression. Grade 12-adult. 1 form (2 forms of 8 tests). 3 (5) — 35 (40) min. per test, (6) hrs. for battery. Test booklets ($3.20 per 25, each); aptitude classification sheet (70¢ per 25); personnel director's booklet, pp. 27 (50¢); counselor's booklet, pp. 35 (40¢); technical supplement, pp. 16 (40¢); student's booklet, pp. 20 (25¢); examiner's manual, pp. 27 (30¢); specimen sets ($3.00-$3.75). Chicago: Science Research Associates, 1953, 1954.

The impressive Flanagan test battery is a product of studies that began in 1946 and are still continuing. The 14 tests were devised to measure job elements based on analyses of the critical behaviors required in a number of occupations. For personnel selection and classification, selected batteries of tests may be administered, as recommended for the job or jobs to be filled. For vocational guidance, the entire six-hour battery may be used to obtain composite scores representing 30 occupations. The technical quality of the tests is excellent. Items have been invented in-

geniously and are contained in well-planned booklets, most of which are arranged for easy self-scoring with carbon inserts. The reliabilities of the test elements range from about .55 to .85, but the reliabilities of the composite scores for occupations, each based on 2 to 7 subtests, are from .83 to .93. Test intercorrelations are generally low, with an average of .29 and only a few coefficients above .50. Norms, based on 1,563 Pittsburgh high school seniors, are expressed as stanines (single-digit standard scores) and as percentiles. Validity for occupational prediction is discussed in terms of a four-year follow-up study of high school students who entered seven occupations; raw correlations with criteria of vocational success range from .29 to .64. The five manuals provide a wealth of data cast in forms appropriate for the various types of users of the tests. In spite of the eight years of intensive work already given to these tests, they still have shortcomings. Some of the defects are easily apparent and are likely to be remedied by continued research, such as the needs for greater empirical validation with respect to specific occupations, and for more broadly based norms. Perhaps the most serious present fault comes from basing all norms on high school students. Thus, an average stanine score of 5 seems to imply average potential for every occupation, whether electrician or physicist, draftsman or engineer, salesperson or psychologist. In addition to the differential ability patterns of these occupations, *levels* must also be taken into account, or counseled students may be seriously misled. Expectancy tables based on the testing of persons in each occupation are needed for fully valid interpretation. Meanwhile, a student should be told, for example, that he might become an acceptable mechanic with a 4 on the mechanic scale, but to become a mathematician at all might require a 9 on the scale appropriate to that occupation. In spite of such remediable defects, the Flanagan tests seem now ready for use by personnel and counseling psychologists who are sufficiently well trained and discriminating. With further development, they are likely to become a very valuable tool.—*L. F. S.*

Books Received

Bauchard, Philippe. *The child audience.* Paris: UNESCO, 1953. Pp. 198 (paper). $2.00.

Bettelheim, Bruno. *Symbolic wounds.* Glencoe, Ill.: Free Press, 1954. Pp. 286. $4.75.

Feingold, Norman. *How to choose that career.* Cambridge, Mass.: Bellman Publishing Co., 1954. Pp. 52 (paper). $1.00.

Frank, Lawrence K., & Frank, Mary. *How to be a woman.* New York: Bobbs-Merrill, 1954. Pp. 144. $2.75.

Gillin, John, *et al. For a science of social man.* New York: Macmillan, 1954. Pp. vii + 289. $4.00.

Gray, J. Stanley. *Psychology applied to human affairs.* New York: McGraw-Hill, 1954. Pp. vii + 581. $6.00.

Ikin, A. Graham. *Life, faith and prayer.* New York: Oxford, 1954. Pp. 127. $2.50.

Lloyd-Jones, Esther, & Smith, Margaret Ruth. (Eds.) *Student personnel work as deeper teaching.* New York: Harper, 1954. Pp. xii + 361. $5.00.

Lorand, Sandor. (Ed.) *The yearbook of psychoanalysis.* New York: International Universities Press, 1954. Pp. 350. $7.50.

McCormick, Mary J. *Diagnostic casework in the Thomistic pattern.* New York: Columbia Univer. Press, 1954. Pp. xii + 239. $3.50.

Moloney, James Clark. *Understanding the Japanese mind.* New York: Philosophical Library, 1954. Pp. xix + 252. $3.50.

Pillay, A. P., & Ellis, Albert. (Eds.) *Sex, society and the individual.* Bombay, India (Whiteaway Bldg.): International Journal of Sexology, 1953. Pp. x + 448. $7.50.

Podolsky, Edward. (Ed.) *Music therapy.* New York: Philosophical Library, 1954. Pp. xii + 355. $6.00.

Sappenfield, Bert R. *Personality dynamics.* New York: Alfred A. Knopf, 1954. Pp. xiv + 412 + xvi. $5.50.

Simpson, George. *Man in society.* New York: Doubleday, 1954. Pp. vi + 90 (paper). 95¢.

Vedder, Clyde B. *The juvenile offender.* Garden City, N. Y.: Doubleday, 1954. Pp. xii + 510. $6.00.

Pastoral Psychology

A monthly devoted to the practical synthesis of the principles and dynamics of clinical psychology, dynamic psychiatry, and psychiatric social work with spiritual and religious values.

EDITORIAL ADVISORY BOARD

REV. OREN H. BAKER
The Colgate-Rochester Divinity School

DANIEL BLAIN, M.D.
Medical Director American Psychiatric Ass'n.

REV. JOHN S. BONNELL
Fifth Avenue Presbyterian Church, N.Y.C.

REV. ERNEST E. BRUDER
St. Elizabeths Hospital Washington, D.C.

REV. ROY A. BURKHART
First Community Church Columbus, Ohio

REV. RUSSELL L. DICKS
The Divinity School Duke University

REV. DAVID D. EITZEN
School of Religion Univ. of So. Calif.

REV. ROLLIN J. FAIRBANKS
Episcopal Theological School, Cambridge

LAWRENCE K. FRANK
New York City

REV. MYRON T. HOPPER, Dean
College of the Bible Lexington, Ky.

REV. PAUL E. JOHNSON
Boston Univ. School of Theology

REV. HALFORD E. LUCCOCK
Yale Univ. Divinity School

ROLLO MAY, PH.D.
Consulting Psychologist New York City

WILLIAM C. MENNINGER, M.D.
Menninger Foundation

JOHN A. P. MILLET, M.D.
American Rehabilitation Committee

REV. WAYNE E. OATES
The Southern Baptist Theological Seminary

REV. DAVID E. ROBERTS
Union Theological Seminary

CARL R. ROGERS, PH.D.
Univ. of Chicago

REV. PAUL TILLICH
Union Theological Seminary

REV. GRANGER E. WESTBERG
Lutheran Advisory Board

REV. CARROLL A. WISE
Garrett Biblical Institute

LUTHER E. WOODWARD, Ph.D
New York State Dept. of Mental Hygiene

REV. SEWARD HILTNER, *Pastoral Consultant*
SIMON DONIGER, PH.D., *Editor*

Here are a few of our important articles:

RELIGIOUS APPLICATIONS OF PSYCHIATRY	Karl A. Menninger, M.D.
RELIGION, PSYCHOTHERAPY, SELFHOOD	Rollo May, Ph.D.
THEOLOGY OF SEX RELATIONS	Reuel Howe, D.D.
THE STRESSES AND STRAINS OF MARRIAGE	Erich Lindemann, M.D.
A BASIC ORIENTATION TO COUNSELING	Carl R. Rogers, Ph.D.
DYNAMICS OF PERSONALITY DEVELOPMENT	Franz Alexander, M.D.
THE UNLOVING PERSONALITY AND THE RELIGION OF LOVE	Bonaro W. Overstreet
THE DIAGNOSTIC USE OF THE BIBLE	Wayne E. Oates, D.D.
THE ROOTS OF RELIGION	Gordon W. Allport, Ph.D.
THE PASTOR AND THE CLINICAL PSYCHOLOGIST	Molly Harrower, Ph.D.
WHAT IS HAPPENING TO THE AMERICAN FAMILY?	Margaret Mead, Ph.D.
THE ART OF DREAM INTERPRETATION	Erich Fromm, Ph.D.
ONE WAY TO PERSONALITY	Lawrence K. Frank
THE ACT OF SURRENDER IN THE TREATMENT OF THE ALCOHOLIC	Harry M. Tiebout, M.D.
GROUP THERAPY IN SEXUAL MALADJUSTMENT	Abraham Stone, M.D.
SOCIAL CONTEXT OF PERSONALITY	Dean Liston Pope, D.D.
SIN AND SALVATION	Dean Willard Sperry, D.D.

PASTORAL PSYCHOLOGY
Great Neck, N. Y.

I enclose — $5 for a one-year subscription $11 for a three-year subscription
(Canadian and foreign subscriptions 50¢ per year additional)

Name .....

Address .....

City..... Zone..... State.....

NEUROLOGY & PSYCHIATRY

Section VIII of Excerpta Medica

The Neurology Section includes abstracts of all articles on neurological surgery in addition to general clinical neuro-physiology.

The section on Psychiatry covers the topics usually contained within the meaning of the word. In addition, the following special branches are dealt with fully:

* **Psychoanalysis, individual and analytical psychology**
* **Sexology, criminology, alcoholism, and drug addiction as related to psychiatry**
* **Social and industrial psychology and psychiatry, vocational guidance and mental testing**
* **Psychometrics**
* **Personality testing and the Rorschach test**
* **Heredity and statistical studies as they apply to psychiatry**
* **Mental defect and epilepsy**

Annual volume of 1000 pages, including an index classified by author and subject.

Subscription price $22.50 per annual volume

EXCERPTA MEDICA SERVICE CORPORATION

250 Madison Ave. **New York 16, N.Y.**

Please enter my subscription to Section VIII, NEUROLOGY AND PSYCHIATRY, OF EXCERPTA MEDICA. I enclose $22.50.

Name.....

Address.....

.....